# Additional (Not-for-Publication) Appendix for: Long-Horizon Stock Valuation and Return Forecasts Conditional on Demographic Projections *

Chaoyi Chen†
Magyar Nemzeti Bank

Nikolay Gospodinov‡
Federal Reserve Bank of Atlanta

Alex Maynard§
University of Guelph

Elena Pesavento¶
Emory University

January 17, 2022

*JEL Classification*: G14, G17, C22.

*Keywords*: demographics, stock market valuation, stock return prediction, conditional forecasts, long-horizon forecasts

*The views expressed here are the authors' and not necessarily those of the Federal Reserve Bank of Atlanta, the Federal Reserve System, or the Magyar Nemzeti Bank (Central Bank of Hungary). Please see the footnote to our main paper for our acknowledgments and research funding.

†Magyar Nemzeti Bank (Central Bank of Hungary), Budapest 0165, Hungary; Email: chenc@mnb.hu.

‡Research Department, Federal Reserve Bank of Atlanta, 1000 Peachtree St. N.E., Atlanta, Georgia, 30309–4470, USA; Email: nikolay.gospodinov@atl.frb.org.

§Department of Economics, University of Guelph, Guelph, Ontario, N1G 2W1, Canada; Email: maynarda@uoguelph.ca.

¶Department of Economics, Emory University, Atlanta, Georgia, 30322–2240, USA; Email: epesave@emory.edu.

## B1 Unit Roots and Cointegration Tests

Table B1 provides unit root test results on the dividend-price ratio ($dp$), middle-old ratio ($MO$), middle-young ratio ($MY$) and the log return with dividends ($r$) using the Augmented Dickey-Fuller Generalized Least Squares (ADF-GLS) test of Elliott et al. (1996). Panel A presents results for the full sample (1901-2015), whereas panel B presents results for the later part of the sample (1947-2015). A maximum of four lags was employed. The row corresponding to the lag with the minimum Modified Aikake Information Criterion (MAIC) value is in bold font and the row corresponding to the minimum Bayesian Information Criterion (BIC) is in italic. The tests allow for both an intercept and a linear trend under the alternative.

The results are somewhat dependent on the lag length and sample period. For both the $dp$ and $MY$ ratios we fail to reject at all but the BIC selected lag-length in the full sample. For $dp$ we also fail to reject at all lag lengths in the later part of the sample, whereas for $MY$ we reject for both the BIC and MAIC lag length choice. The tests for $MO$ reject using both BIC and MAIC in both samples. As expected, we strongly reject a unit root for all cases in the return series.

The failure to reject a unit root in the valuation and, in some cases, the demographic ratios does not necessarily imply a true or exact unit root in these series. The power of unit root tests is well known to be low when the roots are close, but not equal to, unity. Also, there are good a priori reasons to rule out a literal unit root in a ratio variable, which must be bounded between zero and one. Nonetheless, at the very least, the test results confirm that all three variables are highly persistent.

Favero et al. (2011) report evidence of cointegration between the $dp$ and $MY$ ratios. In Table B2, we next test for cointegration among $dp$, $MO$, and $MY$, both in pairs and all together. Panels I and II, respectively, provide results for the Engle-Granger two-step and Johansen cointegration tests. Within each panel, Sub-Panels A and B provide results for the full sample (1901-2015) and post-WWII sample (1947-2015) periods, respectively.

The Engle-Granger two-step cointegration test results at the 5% significance level are presented in the first column of each sub-panel in Panel I. This tests the null hypothesis of no cointegration (unit root in the residuals), with a rejection of this null corresponding to the presence of cointegration. The $p$-value is given in the second column and the AIC value corresponding to each lag choice is given in the third column. The rows corresponding to the minimum AIC value are in bold font. A maximum of five lags was employed. The tests allow for both an intercept and a linear trend under the alternative.

The results are again somewhat sensitive to the lag length and sample. Overall, the Engle-Granger test result does not provide strong evidence of cointegration between the $dp$ ratio and the the two demographic variables. The only case in which we can reject the null of no cointegration at the 5% level at the AIC-selected lag length is for $(dp, MO)$ in the 1947-2015 period. In all other cases, we fail to reject in favor of cointegration at the AIC-selected lag length.

The results of the Engle-Granger tests run somewhat contrary to the earlier findings of Favero et al. (2011) and raises the possibility of a spurious relationship between the financial ratios and demographic variables. However, as in the case of the unit root tests, the failure to reject does not imply acceptance of the null hypothesis, especially when the power of the test is known to be low.[1] Moreover, our specification in linking the dividend-price ratio and demographic ratios in (A1.1) includes a lag of the dividend-price ratio, guaranteeing the stationarity of the residual and thus ruling out the possibility of a spurious regression. As discussed in the paragraph below, the Johansen tests are also generally more supportive of cointegration.

The Johansen test results are provided in the first column of each sub-panel in Panel II and the corresponding $p$-values are given in the second column. This tests the null of only one cointegrating vector against the alternative of one or more cointegrating vectors at the 5% significance level. A linear trend is allowed under the alternative hypothesis. Generally, the Johansen test provides much stronger evidence of cointegration than the Engle-Granger test, with the exception of the $(dp, MY)$ pair over the full sample.

# B2 In-Sample Estimation

## B2.1 Results using the middle-old ratio

The results in Tables B3 and B4 extend the results in Tables 1 and 2 to include the middle-to-old (MO) ratio either in place of, or together with the MY ratio. As can be seen from the two tables, the MO ratio is a less useful demographic predictor than the MY ratio.

[1] The Engle-Granger test is expected to have lower power in the presence of strong endogeneity (Pesavento (2004)).

## B3 Out-of-Sample Forecasts

### B3.1 One Year Pseudo Out-of-Sample Results

Table B5 provides results for one year pseudo out-of-sample. We run recursive forecasts regression for one year hear horizons (k=1). Column 1 lists the candidate models. HM is the historical mean. PR is the classical predictive regression model, using the lagged dividend-price ratio as the predictor. FGT stands for the model in the spirit of Favero et al. (2011), which provides forecasts based on (6) for $k = 1$. MO, MY and MO & MY provide the estimates of (A1.4) using the actual future values of the demographic ratios $MO_{t+1}$, $MY_{t+1}$ and $(MO_{t+1}, MY_{t+1})$ to replace the demographic ratio projections replacing $\widehat{dr}_{t+1|t}$ in (A1.3). The use of the actual future demographic values is appropriate when assessing the effectiveness of the model in providing *conditional* forecasts, where the conditioning is based on the assumption of a correct demographic forecast.[2]

Within the first part of each panel, columns 2,4, and 6 of Table B5 show the out-of-sample mean square errors (OOS MSE) of each model under various training periods. Consistent with the in-sample finding, MY dominates all of the other models in panel A and performs as one of the best models in panel B. Interestingly, even though it places second to MY in most cases, FGT performs consistently better than the remaining models. Indeed, following Favero et al. (2011), our finding gives more evidence that the MY ratio does improve the short-run prediction. Columns 3, 5, and 7 present the out-of-sample $R^2$ (OOS $R^2$), introduced by Campbell and Thompson (2007).[3] The results support the conclusion that MY outperforms the prevailing historical mean during the full sample (Panel A), whereas none of the models show a positive OOS $R^2$ during the post WWII period (Panel B). On the other hand, MY outperforms the predictive regression model with a higher OOS $R^2$ in all cases.

The second part of each panel of Table B5 reports tests of the demographic models against two benchmark forecasts: the historical mean and the predictive regression. The demographic models nest both benchmarks. The Diebold-Mariano (DM) test is known to follow a non-standard distribution when the forecast models are nested (Clark and McCracken (2001) and McCracken

[2]This is because the accuracy of the conditional forecast does not depend on the accuracy of the demographic forecast itself, since we condition on it. By contrast, the unconditional forecast accuracy does depends on the accuracy of the demographic forecast and cannot be accessed in this manner without the introduction of severe look-ahead bias. In Section 4.3, we later assess the accuracy of the unconditional forecast, using the real-time Census Bureau historical forecasts available for $\widehat{dr}_{t+1|t}$. Given the limited availability of the historical Census Bureau forecasts, our primary focus up until Section 4.3 is on the comparison of the models in terms of their ability to provide *conditional* out-of-sample forecasts.

[3]The OOS $R^2$ is defined as $R^2 = 1 - \frac{\text{OOS MSE}}{\text{HM MSE}}$, where $\frac{\text{OOS MSE}}{\text{HM MSE}}$ is the ratio of the out-of-sample mean squared error (OOS MSE) of the model to that of a baseline forecast projecting the historical mean (HM MSE).

(2007)). This results in a bias that inflates the OOS MSE of the larger model under the null hypothesis that the two models provide equal forecasts in large sample. Clark and West (2007) provide a bias correction to the DM test resulting in a $t$ test with a standard normal distribution. It is this Clark and West (CW) correction to the DM test that we employ to examine the significance of the out-performance.[4]

The second part of Columns 2, 4, and 6 show the CW test for out-performance relative to the historical mean under each training period. In five out of six cases, the MY model outperforms the historical mean at the 10% significant level or better. Columns 3, 5, and 7 provide the CW test with the predictive regression as the benchmark model. For the MY model, the test statistics are all significant at the 10% levels and many are also significant at the 5% and 1% levels. This confirms the significance of MY's out-performing OOS MSE and OOS $R^2$ from the first part of Panel A (full sample). In Panel B, the test results appear more favorable to the MY model than the corresponding OOS MSE or OOS $R^2$ results. Indeed the CW test points to out-performance of the MY even for some cases in which its MSE slightly exceeds that of the benchmark. This is due to the bias correction of CW, which adjusts for the fact that the larger nested model – in this case the demographic model – is expected to have the larger OOS MSE even under the null hypothesis in which the two models are equivalent. Therefore, if the MSE of the larger model is even close to matching that of the smaller model it nests, this may be strong enough evidence to reject the null in its favor.

Intriguingly, in many cases, the $P$ values of FGT seem to be fairly close to the $P$ values of MY. Intuitively, this is natural due to the fact that, for one-year ahead forecasting, the only difference between FGT and MY is the replacement of $MY_t$ by $MY_{t+1}$. Considering that the demographic ratio cannot vary greatly in the short-run, it is reasonable to expect $MY_{t+1}$ to contribute only marginally to the information already contained in $MY_t$. As a result, in the short-run, FGT and MY are most likely to show similar forecasting accuracy.

### B3.2 One Year Pseudo Out-of-Sample Rolling Forecasts

To complement our recursive out-of-sample forecasting, we run rolling conditional forecasting regressions for the one year ahead horizon. At the one-year horizon, shown in Table B6, the MY again has the lowest OOS MSE and a positive OOS $R^2$ in five of six cases and nearly ties the historical mean for the lowest $MSE$ in the last case.

[4] For some caveats on applying tests of equal forecast accuracy to conditional forecast models see Clark and McCracken (2017) and Faust and Wright (2008).

As in the main paper, we examine the significance of the out-performance of the rolling window using the Giacomini and White (2006) test, which is designed for rolling windows and applicable to nested models. The second parts of the panels in Table B6 show the GW test for out-performance in rolling samples relative to the historical mean predictive regression. In the short-run, contrary to the recursive results of Table B5, using this more stringent test, $MY$ never significantly outperforms either the historical mean model or the predictive regression model at the one hear horizon, despite its out-performance in terms of OOS MSE and $R^2$. However, we do observe some evidence on the ability of $MY$ to predict long-run stock returns even using the GW test.

To explore the sensitivity of our findings to the training periods and window size, we provide Figures B1 and B2 to show the OOS $R^2$ and p-values for out-performance against the historical mean (vertical axis) for a range of training periods (horizontal axis). Figure B1 show the one-year ahead recursive forecasts, expanding on Tables B5 while Figures B2 and shows the corresponding rolling forecasts and expand on Table B6. The graphical evidence demonstrates the robustness of the results reported earlier in the tables to the choice of training period (recursive) and window size (rolling). The good performance of both the $MY$ and $MY_R$ appears quite robust, with both showing an OOS $R^2$ that remains stable across training periods and window lengths. Moreover, the forecasting results provide evidence that an accurate demographic ratio projection can improve return prediction, particularly in the long run. Consistent with Favero et al. (2011), the results support the use of the middle-to-young (MY) ratio as the best predictor. Moreover, from a conditional forecast perspective, we find that the demographic ratio projection performs better than the FGT model in both the short and long run.

### B3.3 Five Years Results Including MO

Tables B12 and B7 present the same five-years ahead pseudo out-of-sample results that are in the main paper with the additional inclusion of the middle-to-old ratio (MO).

## B4 True Out-of-Sample Forecasts

### B4.1 True Out-of-Sample Forecasts using Historical Census Bureau Predictions

The preceding out-of-sample analysis conditioned on the future projected demographic projection. We now ask whether the same approach can be useful for out-of-sample or real-time forecasts. This requires us to replace the actual future demographic ratios by the corresponding Census Bureau projections that would be available to the forecaster at the time of the prediction. To this end, we collect, compile, and electronically record, historical forecasts from the U.S. Census Bureau. This is complicated by the many years in which the Census Bureau did not update their projections, thereby creating gaps in their historical forecast record. We address these gaps by building our forecasts with the latest available demographic projection that would have been available to the forecaster at the time that the forecast was produced, as described in detail in Appendix A2.

Table B8 presents the five-year ahead out-of-sample forecasts that we have in the main paper. The only difference that we now also include a model that use future demographic ratio for MO and for both MY and MO. As in our earlier tables, these show the results for three different choices of training periods (recursive) or window sizes (rolling). Panel (B) of Table B8, shows that in terms of MSE, the two $MY$-based forecasts out-perform the forecasts obtained using other demographic variables even when using in combination with MO. Overall, these results confirm the predictive content of the Census Bureau $MY$ forecast for long-horizon stock return forecasts.

## B5 Sensitivity to IVX tuning parameters

### B5.1 In-Sample Using IVX

In this section we present the additional sensitivity results we have for in-sample estimation. Tables B9 and B10 present a robustness check on the IVX estimation of Table 1 and 2 of the main paper, by changing the values of IVX tuning parameters to $c = -0.9$, $\alpha$=1. The results show that the qualitative results in Tables 1 and 2 are not driven by the particular IVX tuning parameters employed.

### B5.2 Out-of-Sample Using IVX

Tables B11-B14 and Figures B3-B6 of this Appendix present conditional out-of-sample forecasting results when IVX is used in place of OLS in the first stage forecast for the dividend-price ratio. The results are mainly consistent with those of Table 4 and Figure 3 of our main paper, which provide the equivalent conditional forecast results using OLS instead of IVX in the first step. Notably, for both the one-year and five-year forecasts, the recursive results remain quite strong and similar to those found in the paper using OLS. For the rolling sample forecasts, the use of IVX leads to moderately weaker short-horizon forecast results. However, the merits of using $MY$ as a predictor in the long-run are still clear. For example, the $MY$ based forecasts (either $MY_R$ or $MY_D$) always show the lowest MSE for the five-year forecast. In particular, in the post-WWII sample, $MY$ significantly outperforms both $HM$ and $PR$, and the results are quite robust against the choice of the window size. Since IVX reduces estimation bias at the expense of estimator variance and since the smaller rolling sample sizes entail greater variance, it is understandable that IVX estimation turns out to be more beneficial for recursive than for rolling sample forecasting.

## B6 Additional Out-Of-Sample Forecast Results

In this section of the Appendix we provide results that are supplementary to Section 4.3 of the main paper. Table B15 and Figure B7 show the five-year out-of-sample results when using the real-time method in place of the stale forecast method as described in Section A2 of the Appendix to the main paper. The results are quite similar to those of Table 5 and Figure 4 of the paper. In Table B16 and Figure B8 we provide the unconditional one-year ahead forecast results that would be otherwise equivalent to the five-year ahead forecast results shown in Table 5 and Figure 4. These use the stale forecast method. Similar results using the real-time forecast method are available upon request. As found for many of the conditional results, the advantages of the demographic forecasts are more apparent at the longer five-year horizon than they are at the one-year horizon. This is not surprising given the slowly evolving nature of the demographic ratios.

## B7 Additional COVID19 Forecast Scenarios

In this section we present the results for Scenarios 1-3 of COVID19 scenario analysis from Sections 5.3. The scenarios and model are detailed in Section A3 of the appendix to the main paper. Results for Scenario 4 (our worst case scenario) were included in the paper.

For each of Scenarios 1-3, we compare baseline and COVID19-adjusted projections for the $MY$ ratio, the dividend-price ratio, and stock returns equivalent to those shown in three panels Figure 7 of the main paper for Scenario 4. Figures B9, B10, and B11 provide the results for our best case scenario, Scenario 1 (Quick End, Mild Endemic). Figures B12, B13, and B14 show the results for Scenario 2 (Quick End, Severe Endemic). Finally, Figures B15, B16, and B17 show the results for Scenario 3 (Slow End, Mild Endemic).

### B7.1 COVID19 Scenarios: Details on deaths and removals

For each scenario we plot the actual (2020-2021) and projected deaths due to COVID19 and the cumulative removal (or shortfall) $R_{Y,t}$ and $R_{M,t}$ due to COVID19 deaths. Recall from Section A3.4 of the appendix to the main paper, that the cumulative removals refers to the cumulative impact on the population age range of current and prior year COVID19 fatalities. Put another way, they are the amount that we adjust $Y_t$ and $M_t$ by in order to account for the impact of COVID19. Figures B18 and B19 show death and cumulative removals for Scenario 1. The equivalent Figures for Scenarios 2, 3, and 4 are shown in Figures B20- B21, B22- B23, and B24- B25, respectively.

Table B1: **Augmented Dickey Fuller - Generalized Least Squares Test**

| | | **Panel A: 1901-2015** | | **Panel B: 1947-2015** | |
|---|---|---|---|---|---|
| Time Series | lag length | ADF -GLS | Unit Root | ADF -GLS | Unit Root |
| $dp_t$ | 0 | $-3.75^{***}$ | *Reject* | -2.48 | *Fail* |
| | 1 | -2.78 | Fail | -2.19 | Fail |
| | 2 | -1.96 | **Fail** | -1.82 | **Fail** |
| | 3 | -2.07 | Fail | -2.09 | Fail |
| | 4 | -2.14 | Fail | -2.60 | Fail |
| $\ln(MO_t)$ | 0 | -0.56 | Fail | -0.21 | Fail |
| | 1 | $-4.24^{***}$ | **Reject** | $-3.55^{**}$ | **Reject** |
| | 2 | $-4.94^{***}$ | Reject | $-4.22^{***}$ | Reject |
| | 3 | $-5.04^{***}$ | Reject | $-4.41^{***}$ | Reject |
| | 4 | $-6.22^{***}$ | *Reject* | $-5.72^{***}$ | *Reject* |
| $\ln(MY_t)$ | 0 | -0.91 | Fail | -0.60 | Fail |
| | 1 | -1.67 | **Fail** | -1.94 | Fail |
| | 2 | -1.99 | Fail | $-3.23^{**}$ | **Reject** |
| | 3 | -2.45 | Fail | $-4.30^{***}$ | Reject |
| | 4 | $-2.91^{*}$ | *Reject* | $-5.40^{***}$ | *Reject* |
| $r_t$ | 0 | $-10.59^{***}$ | **Reject** | $-8.45^{***}$ | **Reject** |
| | 1 | $-9.24^{***}$ | *Reject* | $-7.13^{***}$ | Reject |
| | 2 | $-6.38^{***}$ | Reject | $-5.05^{***}$ | Reject |
| | 3 | $-5.97^{***}$ | Reject | $-3.80^{***}$ | Reject |
| | 4 | $-6.23^{***}$ | Reject | $-2.98^{*}$ | Reject |

The table shows the results from the Augmented Dickey Fuller Generalized Least Squares (ADF-GLS) unit root test of Elliott et al. (1996) for the four time series shown in the first column. Column 2 (lag length) shows the number of lagged first differences included in the ADF-GLS regression specification. Results for the time period 1901-2015 are shown in Panel A (Columns 3-4) and results for 1947-2015 period are shown in Panel B (Columns 5-6). Column 3 and 5 show the value of the ADF-GLS test statistic, with one, two and three stars(*) denoting rejection at 10%, 5% and 1% respectively. Column 4 and 6 denotes whether the unit root was rejected at least at the 10% level, with "reject" noting a rejection of the unit root hypothesis, and "Fail" noting a failure to reject. All the tests are run with an intercept and time trend. The critical values for the ADF-GLS test are -3.70 (1%), -3.13 (5%) and -2.83 (10%). Bold indicates the lag length chosen by the MAIC criteria with a maximum of 4 lags. Italics indicates the lag length chosen by BIC.

Table B2: **Cointegration Tests**

| | | Panel-I: Engle-Granger 2-Step | | | | | | Panel -II: Johansen | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Sub-Panel A: 1901-2015 | | | Sub-Panel B: 1947-2015 | | | Sub-Panel A: 1901-2015 | | Sub-Panel B: 1947-2015 | |
| Time Series | lag length | No Cointegration | P Value | AIC Value | No Cointegration | P Value | AIC Value | No Cointegration | P Value | No Cointegration | P Value |
| dp | 0 | Reject | 0.045 | -22.88 | **Reject** | **0.034** | **-56.245** | Fail | 0.549 | Fail | 0.733 |
| & | 1 | Fail | 0.081 | -19.949 | Fail | 0.077 | -53.769 | Reject | 0.002 | Reject | 0.03 |
| MO | 2 | Fail | 0.415 | -24.93 | Fail | 0.29 | -56.207 | Reject | 0.004 | Reject | 0.007 |
| | 3 | **Fail** | **0.258** | **-25.616** | Fail | 0.152 | -55.307 | Reject | 0.002 | Reject | 0.001 |
| | 4 | Fail | 0.214 | -22.988 | Fail | 0.075 | -54.974 | Reject | 0. 001 | Reject | 0.001 |
| | 5 | Fail | 0.284 | -19.929 | Fail | 0.136 | -50.814 | Reject | 0.016 | Reject | 0.003 |
| dp | 0 | Reject | 0.004 | -35.663 | Reject | 0.023 | -56.182 | Reject | 0.02 | Fail | 0.56 |
| & | 1 | Reject | 0.011 | -33.239 | Fail | 0.069 | -53.1 | Fail | 0.09 | Reject | 0.027 |
| MY | 2 | Fail | 0.22 | -42.447 | **Fail** | **0.404** | **-56.291** | Fail | 0.419 | Reject | 0.01 |
| | 3 | **Fail** | **0.136** | **-43.761** | Fail | 0.37 | -55.072 | Fail | 0.269 | Reject | 0.008 |
| | 4 | Fail | 0.115 | -42.025 | Fail | 0.2 | -53.517 | Fail | 0.175 | Reject | 0.001 |
| | 5 | Fail | 0.14 | -41.23 | Fail | 0.302 | -49.544 | Fail | 0.096 | Reject | 0.001 |
| dp, | 0 | Reject | 0.005 | -34.52 | Fail | 0.056 | -34.52 | Reject | 0.04 | Reject | 0.02 |
| MO | 1 | Reject | 0.0150 | -32.1018 | **Fail** | **0.151** | **-52.766** | Reject | 0.03 | Reject | 0.02 |
| & | 2 | Fail | 0.254 | -41.448 | Fail | 0.596 | -49.549 | Reject | 0.04 | Reject | 0.01 |
| MY | 3 | **Fail** | **0.154** | **-42.668** | Fail | 0.596 | -52.649 | Reject | 0.02 | Reject | 0.004 |
| | 4 | Fail | 0.131 | -40.836 | Fail | 0.404 | -51.027 | Reject | 0.001 | Reject | 0.001 |
| | 5 | Fail | 0.159 | -39.9 | Fail | 0.5 | -45.639 | Reject | 0.003 | Reject | 0.001 |

The table shows both the two-step Engle-Granger and the Johansen cointegration tests of the null hypothesis of no cointegration. The combination of variables tested are shown in Column 1, where dp is the log dividend-price ratio, and MO and MY are the middle-to-old and middle-to-young ratios, respectively. Column 2 (lag length) shows the number of lagged first differences included in the Engle-Granger specification. Engle-Granger test results are provided in Panel I and Johansen test results are provided in Panel II. Within each panel results for the time period 1901-2015 are shown in Sub-Panel A (Columns 3-5) and results for 1947-2015 period are shown in Sub-Panel B (Columns 6-8). Columns 3 and 6 (No cointegration) the result of testing the null hypothesis of no cointegration at the 5% significance level, with "reject" noting a rejection of the unit root hypothesis, and "Fail" noting a failure to reject. The Engle-Granger test P-values are provided in Columns 4 and 7. The Aikake Information Criterion (AIC) for the number of lagged differences is shown in Columns 5 and 8. The row with the lag-length that minimizes this criteria is shown in bold. The result at the 5% level for the Johansen test of the null of no cointegrating vectors, with the alternative of one or more cointegrating vectors are shown in Columns 9 and 11. The corresponding p-values are shown in Columns 10 and 12.

Table B3: **Return Regression Models: OLS and IVX Estimation Results**

| Model: | lag dp | | MO and lag dp | | MY and lag dp | | MO, MY, and lag dp | |
|---|---|---|---|---|---|---|---|---|
| | OLS | IVX | OLS | IVX | OLS | IVX | OLS | IVX |
| | | | | **Panel A: 1901-2015** | | | | |
| Const | 0.257*** | 0.257** | -0.065 | -0.065 | -0.022 | -0.022 | -0.025 | -0.025 |
| $\widehat{dp}_{t+1\|t}$ | | | -0.8736 | -0.418 | $-0.755$*** | $-0.649$** | $-0.765$** | $-0.691$** |
| $dp_t$ | 0.062 | 0.062 | 0.839* | 0.434 | 0.733** | 0.642** | 0.742*** | 0.681** |
| Model Test | 2.615 | 2.537 | 1.857 | 2.831 | 5.789*** | 7.817** | 6.343*** | 8.195** |
| p-value | 0.109 | 0.111 | 0.161 | 0.243 | 0.004 | 0.020 | 0.003 | 0.017 |
| | | | | **Panel B: 1947-2015** | | | | |
| Const | 0.402** | 0.402** | 0.097 | 0.097 | 0.032 | 0.032 | 0.028 | 0.028 |
| $\widehat{dp}_{t+1\|t}$ | | | -0.987 | -1.742 | $-1.198$*** | $-1.549$*** | $-1.210$*** | $-1.534$*** |
| $dp_t$ | 0.098** | 0.114* | 1.001* | 1.732 | 1.194*** | 1.550*** | 1.205*** | 1.535*** |
| Model test | 5.008** | 3.561* | 3.780** | 3.980 | 6.509*** | 12.250*** | 6.610*** | 12.807*** |
| p-value | 0.029 | 0.0589 | 0.028 | 0.1367 | 0.003 | 0.0022 | 0.002 | 0.0017 |

$***$ significantly different from zero at the 1% level, $**$, significantly different from zero at the 5% level, $*$ significantly different from zero at the 10% level. This table provides IVX estimation of equations (1) for $k = 1$, (3), and (4). The dependent variable in all cases are yearly log returns including dividends. Column 2 (lag dp), provides the estimates of (1) for $k = 1$ in which only on the past $dp_t$ is employed as a predictor. Columns 3-5 provide estimates of (4), using three different specifications of the demographic ratio projection $dp_{t+1|t}$ in (A1.1): only MO (Column 3), only MY (Column 4), and both MO and MY (Column 5). Results for the time period 1901-2015 are shown in Panel A and results for 1947-2015 period are shown in Panel B. Sensitivity results with respect to these IVX tuning parameters are included in the additional (not-for-publication) appendix.

Table B4: **Dividend-Price Ratio Model: OLS and IVX Estimation Results**

| Model: | AR(1) | | Augmented AR(1) | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | MO and | | MY and | | MO, MY | |
| | lag dp | | lag dp | | lag dp | | and lag dp | |
| | OLS | IVX | OLS | IVX | OLS | IVX | OLS | IVX |
| | | | | **Panel A: 1901-2015** | | | | |
| Const | $-0.368^{**}$ | $-0.368^{**}$ | $-0.513^{**}$ | $-0.513^{**}$ | $-0.395^{***}$ | $-0.395^{***}$ | $-0.529^{***}$ | $-0.529^{***}$ |
| lap dp | $0.889^{***}$ | $0.886^{***}$ | $0.875^{***}$ | $0.887^{***}$ | $0.757^{***}$ | $0.776^{***}$ | $0.745^{***}$ | $0.779^{***}$ |
| MO | | | 0.053 | -0.002 | | | 0.049 | -0.006 |
| MY | | | | | $-0.504^{***}$ | $-0.400^{*}$ | $-0.500^{***}$ | $-0.395^{*}$ |
| Model Test | $395.415^{***}$ | $374.89^{***}$ | $198.339^{***}$ | $384.65^{***}$ | $222.308^{***}$ | $409.87^{***}$ | $148.506^{***}$ | $425.03^{***}$ |
| p-value | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| | | | | **Panel B: 1947-2015** | | | | |
| Const. | $-0.309^{**}$ | $-0.309^{**}$ | -0.227 | -0.227 | $-0.335^{***}$ | $-0.335^{***}$ | $-0.345^{**}$ | $-0.345^{**}$ |
| lap dp | $0.915^{***}$ | $0.904^{***}$ | $0.879^{***}$ | $0.817^{***}$ | $0.829^{***}$ | $0.777^{***}$ | $0.828^{***}$ | $0.778^{***}$ |
| MO | | | -0.130 | -0.227 | | | 0.014 | -0.122 |
| MY | | | | | $-0.312^{**}$ | $-0.408^{**}$ | -0.3297 | -0.254 |
| Model Test | $377.506^{***}$ | $170.61^{***}$ | $193.103^{***}$ | $269.69^{***}$ | $200.331^{***}$ | $228.78^{***}$ | $131.526^{***}$ | $282.94^{***}$ |
| p-value | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

$***$ significantly different from zero at the 1% level, $**$, significantly different from zero at the 5% level, $*$ significantly different from zero at the 10% level. This table provides IVX estimates of (2) & (A1.1). Column 2 provides estimates for the pure AR(1) process in (2). Column 3-5 provide estimates for the augmented AR(1) process including demographic ratios as in (A1.1). In columns 3, 4, and 5 respectively, MO, MY, and both MO and MY are employed respectively as demographic controls. Results for the time period 1901-2015 are shown in Panel A and results for 1947-2015 period are shown in Panel B. For the IVX estimation, we set $c = -1$, $\alpha = 0.95$. Sensitivity results with respect to these IVX tuning parameters are included in the additional (not-for-publication) appendix.

Table B5: **Results of Return Regression Models (Pseudo Out-of-Sample, Recursive, One-Year Ahead Forecast)**

**Panel A: 1901-2015: using an initial training period of $tp$ years.**

| training period ($tp$) | $tp$=30 | | $tp$=40 | | $tp$=60 | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ |
| HM | 0.0354 | 0 | 0.0266 | 0 | 0.0262 | 0 |
| PR | 0.0367 | -0.0368 | 0.0269 | -0.0132 | 0.0274 | -0.0482 |
| FGT | 0.0355 | -0.0020 | 0.0245 | 0.0780 | 0.0254 | 0.0280 |
| MO | 0.0402 | -0.1347 | 0.0275 | -0.0325 | 0.0294 | -0.1237 |
| MY | 0.0347 | 0.0204 | 0.0237 | 0.1090 | 0.0250 | 0.0457 |
| MY & MO | 0.0435 | -0.2274 | 0.0250 | 0.0596 | 0.0261 | 0.0025 |
| Model | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) |
| PR | 0.2071 | | 0.0941* | | 0.2090 | |
| FGT | 0.1029 | 0.0228** | 0.0104** | 0.0035*** | 0.0635* | 0.0216** |
| MO | 0.6797 | 0.6480 | 0.1307 | 0.1784 | 0.4325 | 0.5595 |
| MY | 0.0512* | 0.0077*** | 0.0048** | 0.0014*** | 0.0531* | 0.0156** |
| MY & MO | 0.4416 | 0.2981 | 0.0217** | 0.0094*** | 0.1068 | 0.0569* |

**Panel B: 1947-2015: using an initial training period of $tp$ years.**

| training period ($tp$) | $tp$=20 | | $tp$=25 | | $tp$=30 | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | $R^2$ | OOS MSE | $R^2$ | OOS MSE | $R^2$ |
| HM | 0.0282 | 0 | 0.0293 | 0 | 0.0240 | 0 |
| PR | 0.0292 | -0.0350 | 0.0313 | -0.0666 | 0.0288 | -0.2010 |
| FGT | 0.0306 | -0.0850 | 0.0309 | -0.0536 | 0.0269 | -0.1214 |
| MO | 0.0326 | -0.1582 | 0.0336 | -0.1443 | 0.0283 | -0.1797 |
| MY | 0.0287 | -0.0178 | 0.0295 | -0.0063 | 0.0257 | -0.0726 |
| MY & MO | 0.0326 | -0.1550 | 0.0335 | -0.1403 | 0.0290 | -0.2100 |
| Model | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) |
| PR | 0.1028 | | 0.1654 | | 0.4126 | |
| FGT | 0.0290** | 0.0343** | 0.0859* | 0.0486** | 0.1706 | 0.0179** |
| MO | 0.3432 | 0.2520 | 0.2997 | 0.1826 | 0.3115 | 0.0857* |
| MY | 0.0314** | 0.0275** | 0.0818* | 0.0299** | 0.1602 | 0.0075*** |
| MY & MO | 0.2204 | 0.5535 | 0.3020 | 0.4196 | 0.3833 | 0.1923 |

$***$, $**$, $*$ significantly out-performs the benchmark forecasts at the 1% level, 5% level, and 10% level, respectively. Reported p-values are one-sided. This table provides one-year out-of-sample forecasting results. The dependent variable in all cases are yearly log returns including dividends. Results for the time period 1901-2015 are shown in Panel A and results for 1947-2015 period are shown in Panel B. Column 2-3 reports the out-of-sample mean square error (OOS MSE) and out-of-sample $R^2$ (OOS $R^2$) with 30/20 years training period for panel A/panel B. Column 4-5 shows OOS MSE and OOS $R^2$ with 40/25 years training period for panel A/panel B. Column 6-7 gives OOS MSE and OOS $R^2$ with 60/30 years training period for panel A/panel B. This table also provides the out-performance test results. HM is the out-of-sample historical mean. PR is the predictive regression model. FGT provides the estimates of (A1.4) in which the $\hat{x}_{t+1}$ is estimated by $x_t$ and $MY_t$. MO,MY , and MYMO provide estimates of (A1.4) using three different specifications of the demographic ratio projection $dr_{t+1|t}$ in (A1.3). The OOS $R^2$ refers to the out-of-sample $R^2$ using HM as the benchmark. CW is the Clark and West (2007) test. The columns marked (HM, CW) and (PR, CW) provide p-values for the CW test using HM and PR, respectively, as benchmarks.

Table B6: **Results of Return Regression Models (Pseudo Out-of-Sample, Rolling, One-Year Ahead Forecast)**

**Panel A: 1901-2015: using a rolling window of length $w$.**

| window ($w$) | $w = 30$ | | $w = 40$ | | $w = 60$ | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ |
| HM | 0.0366 | 0 | 0.0272 | 0 | 0.0268 | 0 |
| PR | 0.0375 | -0.0266 | 0.0270 | 0.0057 | 0.0279 | -0.0389 |
| FGT | 0.0380 | -0.0402 | 0.0273 | -0.0071 | 0.0271 | -0.0123 |
| MO | 0.0404 | -0.1056 | 0.0305 | -0.1244 | 0.0333 | -0.2425 |
| MY | 0.0356 | 0.0260 | 0.0268 | 0.0132 | 0.0268 | -0.0005 |
| MY & MO | 0.0478 | -0.3068 | 0.0295 | -0.0852 | 0.0331 | -0.2339 |
| Model | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) |
| PR | 0.6870 | | 0.4625 | | 0.6330 | |
| FGT | 0.6593 | 0.5592 | 0.5228 | 0.5472 | 0.5352 | 0.3901 |
| MO | 0.8476 | 0.7679 | 0.7838 | 0.8034 | 0.9384 | 0.9417 |
| MY | 0.3947 | 0.2801 | 0.4610 | 0.4751 | 0.5013 | 0.3610 |
| MY & MO | 0.9112 | 0.8826 | 0.7068 | 0.7252 | 0.8869 | 0.9187 |

**Panel B: 1947-2015: using a rolling window of length $w$**

| window ($w$) | $w = 20$ | | $w = 25$ | | $w = 30$ | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ |
| HM | 0.0296 | 0 | 0.0308 | 0 | 0.0252 | 0 |
| PR | 0.0287 | 0.0296 | 0.0339 | -0.1013 | 0.0291 | -0.1537 |
| FGT | 0.0283 | 0.0450 | 0.0294 | 0.0473 | 0.0263 | -0.0430 |
| MO | 0.0313 | -0.0587 | 0.0322 | -0.0466 | 0.0271 | -0.0.0731 |
| MY | 0.0278 | 0.0610 | 0.0289 | 0.0630 | 0.0235 | 0.0668 |
| MY & MO | 0.0410 | -0.3839 | 0.0389 | -0.2621 | 0.0306 | -0.2130 |
| Model | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) |
| PR | 0.3961 | | 0.8211 | | 0.9610 | |
| FGT | 0.3610 | 0.4516 | 0.3665 | 0.1756 | 0.6482 | 0.2326 |
| MO | 0.6194 | 0.6666 | 5979 | 0.3973 | 0.6138 | 0.3788 |
| MY | 0.2931 | 0.3892 | 0.3060 | 0.1292 | 0.2632 | 0.0598* |
| MY & MO | 0.9258 | 0.9424 | 0.9493 | 0.8080 | 0.8105 | 0.5942 |

$***$, $**$, $*$ significantly out-performs the benchmark forecasts at the 1% level, 5% level, and 10% level, respectively. Reported p-values are one-sided. This table provides one-year out-of-sample forecasting results based on the rolling forecast method. The dependent variable in all cases are yearly log returns including dividends. Results for the time period 1901-2015 are shown in Panel A and results for 1947-2015 period are shown in Panel B. Column 2-3 reports the out-of-sample mean square error (OOS MSE) and out-of-sample $R^2$ (OOS $R^2$) with a 30/20-year training period for panel A/panel B. Column 4-5 shows OOS MSE and OOS $R^2$ with a 40/25-year training period for panel A/panel B. Column 6-7 gives OOS MSE and OOS $R^2$ with a 60/30-year training period for panel A/panel B. This table also provides the out-performance test results. HM is the out-of-sample historical mean. PR is the predictive regression model. FGT provides the estimates of (A1.4) in which the $\hat{x}_{t+1}$ is estimated by $x_t$and $MY_t$. MO,MY, and MYMO provide estimates of (A1.4) using three different specifications of the demographic ratio projection $dr_{t+1|t}$ in (A1.3).The OOS $R^2$ refers to the out-of-sample $R^2$ using HM as the benchmark. GW is the adjusted one-sided Giacomini and White (2006) test. The columns marked (HM, GW) and (PR, GW) provide p-values for the GW test using HM and PR, respectively, as benchmarks.

Table B7: **Results of Return Regression Models (Pseudo Out-of-Sample, Recursive, Five-Year Ahead Forecast)**

**Panel A: 1901-2015: using an initial training period of $tp$ years.**

| training period ($tp$) | $tp = 30$ | | $tp = 40$ | | $tp = 60$ | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ |
| HM | 0.1438 | 0 | 0.1384 | 0 | 0.1171 | 0 |
| PR | 0.1336 | 0.0708 | 0.1333 | 0.0370 | 0.1175 | -0.0037 |
| FGT | 0.1278 | 0.1110 | 0.1239 | 0.1047 | 0.1366 | -0.1668 |
| MO | 0.2123 | -0.4767 | 0.1484 | -0.0723 | 0.1719 | -0.4683 |
| $MY_D$ | 0.1429 | 0.0058 | 0.1279 | 0.0757 | 0.0786 | 0.3287 |
| $MY_R$ | 0.0997 | 0.3065 | 0.0876 | 0.3672 | 0.0881 | 0.2471 |
| MY &MO | 0.5542 | -2.8545 | 0.1212 | 0.1242 | 0.0815 | 0.3043 |
| Model | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) |
| PR | 0.077* | | 0.1186 | | 0.1744 | |
| FGT | 0.0253** | 0.0495** | 0.0518* | 0.0847* | 0.1480 | 0.3101 |
| MO | 0.5195 | 0.6172 | 0.2101 | 0.2609 | 0.6015 | 0.8223 |
| $MY_D$ | 0.0063*** | 0.0075*** | 0.0142** | 0.0147** | 0.0560* | 0.0658* |
| $MY_R$ | 0.0291** | 0.00258** | 0.0320** | 0.0210** | 0.0856* | 0.0588* |
| MY &MO | 0.7268 | 0.7354 | 0.1053 | 0.0565* | 0.0991* | 0.0279** |

**Panel B: 1947-2015: using an initial training period of $tp$ years.**

| training period ($tp$) | $tp = 20$ | | $tp = 25$ | | $tp = 30$ | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ |
| HM | 0.1569 | 0 | 0.1286 | 0 | 0.1168 | 0 |
| PR | 0.1611 | -0.0269 | 0.1634 | -0.2704 | 0.1769 | -0.5148 |
| FGT | 0.1649 | -0.0511 | 0.1709 | -0.3293 | 0.1875 | -0.6058 |
| MO | 0.1717 | -0.0946 | 0.1738 | -0.3513 | 0.1815 | -0.5541 |
| $MY_D$ | 0.0942 | 0.3993 | 0.0788 | 0.3870 | 0.0856 | 0.2670 |
| $MY_R$ | 0.0990 | 0.3689 | 0.1079 | 0.1613 | 0.1200 | -0.0275 |
| MY & MO | 0.2057 | -0.3112 | 0.2057 | -0.5994 | 0.2234 | -0.9133 |
| Model | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) |
| PR | 0.1178 | | 0.1697 | | 0.2675 | |
| FGT | 0.0937* | 0.1418 | 0.1156 | 0.1514 | 0.2004 | 0.2003 |
| MO | 0.0990* | 0.1567 | 0.1123 | 0.1770 | 0.2059 | 0.2405 |
| $MY_D$ | 0.0532* | 0.0785* | 0.0521* | 0.0729* | 0.0810* | 0.0919* |
| $MY$ | 0.0826* | 0.0415** | 0.0895* | 0.0570* | 0.1632 | 0.0758* |
| MY & MO | 0.2168 | 0.9267 | 0.3075 | 0.8996 | 0.4336 | 0.8900 |

$***$, $**$, $*$ significantly out-performs the benchmark forecasts at the 1% level, 5% level, and 10% level, respectively. Reported p-values are one-sided. This table provides five-year out-of-sample forecasting results. The dependent variable in all cases are yearly log returns including dividends. Results for the time period 1901-2015 are shown in Panel A and results for 1947-2015 period are shown in Panel B. The dependent variable in all cases are yearly log returns including dividends. HM stands for historical mean. PR denotes the five-year ahead forecast obtained by forward recursion from the one-year ahead predictive regression forecast. $MY_D$ shows the results from (A1.8) when using MY as the demographic variable with $k = 5$. FGT provides the estimates of (A1.4) in which the $\hat{x}_{t+1}$ is estimated by $x_t$ and $MY_t$. MO, $MY_R$ , and MYMO provide estimates of (A1.4) using three different specifications of the demographic ratio projection $dr_{t+h+1|t}$ in (A1.3) with $h = 4$. The OOS $R^2$ uses the HM as its benchmark and is defined in Footnote 11. CW is the Clark and West (2007) test. The columns marked (HM, CW) and (PR, CW) provide p-values for the CW test using HM and PR, respectively, as benchmarks.

Table B8: **Results of Return Regression Models (True Out-of-Sample, Five-Year Ahead Forecast, Stale)**

**Panel A: Recursive, 1951-2015: using an initial training period of $tp$ years.**

| training period ($tp$) | $tp$=20 | | $tp$=25 | | $tp$=30 | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ |
| HM | 0.1270 | 0 | 0.1157 | 0 | 0.1273 | 0 |
| PR | 0.1531 | -0.2048 | 0.1629 | -0.4074 | 0.1684 | -0.3233 |
| FGT | 0.1784 | -0.4043 | 0.1871 | -0.6169 | 0.2036 | -0.6000 |
| MO | 0.1720 | -0.3538 | 0.1822 | -0.5745 | 0.2060 | -0.6188 |
| $MY_D$ | **0.0945** | 0.2563 | **0.1034** | 0.1063 | **0.1163** | 0.0862 |
| $MY_R$ | 0.1020 | 0.1974 | 0.1087 | 0.0611 | 0.1216 | 0.0443 |
| MY & MO | 0.2242 | -0.7650 | 0.2103 | -0.8175 | 0.2376 | -0.8667 |
| Model | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) |
| PR | 0.1996 | | 0.3341 | | 0.3110 | |
| FGT | 0.1395 | 0.1341 | 0.2681 | 0.1931 | 0.3625 | 0.4877 |
| MO | 0.1233 | 0.1969 | 0.2340 | 0.2770 | 0.2649 | 0.3677 |
| $MY_D$ | 0.0525* | 0.0833* | 0.0844* | 0.1151 | 0.1083 | 0.1665 |
| $MY_R$ | 0.0811* | 0.0563* | 0.1371 | 0.0854* | 0.1730 | 0.1475 |
| MY & MO | 0.4187 | 0.4936 | 0.3492 | 0.4178 | 0.3562 | 0.5123 |

**Panel B: Rolling, 1951-2015: using an initial training period of $tp$ years.**

| training period ($tp$) | $tp$=20 | | $tp$=25 | | $tp$=30 | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | $R^2$ | OOS MSE | $R^2$ | OOS MSE | $R^2$ |
| HM | 0.1754 | 0 | 0.1593 | 0 | 0.1550 | 0 |
| PR | 0.2414 | -0.3761 | 0.2697 | -0.6934 | 0.2137 | -0.3786 |
| FGT | 0.2309 | -0.3164 | 0.3228 | -1.0265 | 0.2653 | -0.7116 |
| MO | 0.2160 | -0.2313 | 0.1702 | -0.0688 | 0.2078 | -0.3404 |
| $MY_D$ | 0.1834 | -0.0453 | **0.1000** | 0.3723 | **0.0915** | 0.4097 |
| $MY_R$ | **0.1009** | 0.4251 | 0.1075 | 0.3251 | 0.1219 | 0.2140 |
| MY & MO | 0.2701 | -0.5395 | 0.1945 | -0.2213 | 0.2035 | -0.3126 |
| Model for $\widehat{dp}_{t+1\|t}$ | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) |
| PR | 0.7981 | | 0.9099 | | 0.8639 | |
| FGT | 0.9163 | 0.4489 | 0.9822 | 0.8576 | 0.9807 | 0.8920 |
| MO | 0.6295 | 0.4387 | 0.5462 | 0.2527 | 0.6689 | 0.4841 |
| $MY_D$ | 0.5803 | 0.2738 | 0.0155** | 0.0481** | 0.0657* | 0.0841* |
| $MY_R$ | 0.0018*** | 0.0546* | 0.0100*** | 0.0455** | 0.0775* | 0.0769* |
| MY & MO | 0.7927 | 0.5728 | 0.6202 | 0.3289 | 0.6523 | 0.4730 |

$***$, $**$, $*$ significantly out-performs the benchmark forecasts at the 1% level, 5% level, and 10% level, respectively. Reported p-values are one-sided. This table provides five-year out-of-sample unconditional forecasting results using the stale forecasts. The dependent variable in all cases are yearly log returns including dividends. Results of recursive method are shown in Panel A and results of rolling method are shown in Panel B. Column 2-3 reports the out-of-sample mean square error (OOS MSE) and out-of-sample $R^2$ (OOS $R^2$) with a 20-year training period for panel A/panel B. Column 4-5 shows OOS MSE and OOS $R^2$ with a 25-year training period for panel A/panel B. Column 6-7 gives OOS MSE and OOS $R^2$ with a 30-year training period for panel A/panel B. This table also provides the out-performance test results. HM is the out-of-sample historical mean. PR is the predictive regression model. FGT provides the estimates of (A1.4) in which the $\hat{x}_{t+1}$ is estimated by $x_t$ and $MY_t$. MO,MY , and MYMO provide estimates of (A1.4) using three different specifications of the demographic ratio projection $dr_{t+1|t}$ in (A1.3). The OOS $R^2$ uses the HM as its benchmark and is defined in Footnote 11. CW is the Clark and West (2007) test. GW is the adjusted one-sided Giacomini and White (2006) test. CW is the Clark and West (2007) test. The columns marked (HM, CW) and (PR, CW) provide p-values for the CW test using HM and PR, respectively, as benchmarks. Similarly, columns marked (HM, GW) and (PR, GW) provide p-values for the GW test.

Table B9: **Dividend-Price Ratio Model IVX Estimation c=-0.9, $\alpha$=1**

| Model: | AR(1) | Augmented AR(1) | | |
|---|---|---|---|---|
| | | MO and | MY and | MO, MY, |
| | lag dp | lag dp | lag dp | and lag dp |
| **Panel A: 1901-2015** | | | | |
| Constant | $-0.368^{**}$ | $-0.513^{**}$ | $-0.395^{***}$ | $-0.529^{***}$ |
| lap dp | $0.888^{***}$ | $0.888^{***}$ | $0.778^{***}$ | $0.770^{***}$ |
| MO | | 0.024 | | 0.010 |
| MY | | | $-0.418^{**}$ | $-0.424^{*}$ |
| Overall test of significance | $382.76^{***}$ | $388.12^{***}$ | $417.08^{***}$ | $429.07^{***}$ |
| p-value | 0 | 0 | 0 | 0 |
| **Panel B: 1947-2015** | | | | |
| Constant | $-0.309^{**}$ | -0.227 | $-0.335^{***}$ | $-0.345^{**}$ |
| lap dp | $0.906^{***}$ | $0.834^{***}$ | $0.792^{***}$ | $0.794^{***}$ |
| MO | | -0.202 | | -0.084 |
| MY | | | $-0.380^{**}$ | -0.274 |
| Overall test of significance | $169.47^{***}$ | $276.74^{***}$ | $229.30^{***}$ | $290.95^{***}$ |
| p-value | 0 | 0 | 0 | 0 |

$***$ significantly different from zero at the 1% level, $**$, significantly different from zero at the 5% level, $*$ significantly different from zero at the 10 % level. This table provides IVX estimates of (2) & (A1.1). Column 2 provides estimates for the pure AR(1) process in (2). Column 3-5 provide estimates for the augmented AR(1) process including demographic ratios as in (A1.1). In columns 3, 4, and 5 respectively, MO, MY, and both MO and MY are employed respectively as demographic controls. Results for the time period 1901-2015 are shown in Panel A and results for 1947-2015 period are shown in Panel B.

Table B10: **IVX Estimation of Return Regression Models (In-Sample), c=-0.9, $\alpha$=1**

| Model: | AR(1) | Augmented AR(1) | | |
|---|---|---|---|---|
| | | MO and | MY and | MO, MY, |
| | lag dp | lag dp | lag dp | and lag dp |
| **Panel A: 1901-2015** | | | | |
| Constant | 0.257** | -0.065 | -0.022 | -0.025 |
| $\widehat{dp}_{t+1\|t}$ | | 0.541 | 0.658** | 0.690** |
| $dp_t$ | 0.061 | -0.540 | $-0.667^{**}$ | $-0.702^{**}$ |
| Overall test of significance | 2.535 | 2.924 | 7.899** | 8.212** |
| p-value | 0.111 | 0.232 | 0.019 | 0.017 |
| **Panel B: 1947-2015** | | | | |
| Constant | 0.402** | 0.097 | 0.032 | 0.028 |
| $\widehat{dp}_{t+1\|t}$ | | 2.209 | 1.456*** | 1.446*** |
| $dp_t$ | 0.110* | -1.580 | $-1.456^{***}$ | $-1.446^{***}$ |
| Overall test of significance | 3.336* | 3.534 | 11.096*** | 11.668*** |
| p-value | 0.068 | 0.171 | 0.004 | 0.003 |

$***$ significantly different from zero at the 1% level, $**$, significantly different from zero at the 5% level, $*$ significantly different from zero at the 10 % level. This table provides IVX estimation of equations (1) for $k = 1$, (3), and (4). The dependent variable in all cases are yearly log returns including dividends. Column 2 (lag dp), provides the estimates of (1) for $k = 1$ in which only on the past $dp_t$ is employed as a predictor. Columns 3-5 provide estimates of (4), using three different specifications of the demographic ratio projection $dp_{t+1|t}$ in (A1.1): only MO (Column 3), only MY (Column 4), and both MO and MY (Column 5). Results for the time period 1901-2015 are shown in Panel A and results for 1947-2015 period are shown in Panel B.

Table B11: **Results of Return Regression Models (Pseudo Out-of-Sample, Recursive, One-Year Ahead Forecast, IVX Estimation)**

**Panel A: 1901-2015: using an initial training period of $tp$ years.**

| training period ($tp$) | $tp$=30 | | $tp$=40 | | $tp$=60 | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ |
| HM | 0.0354 | 0 | 0.0266 | 0 | 0.0262 | 0 |
| PR | 0.0392 | -0.1063 | 0.0268 | -0.0075 | 0.0284 | -0.0848 |
| FGT | 0.0340 | 0.0395 | 0.0237 | 0.1073 | 0.0250 | 0.0464 |
| MO | 0.0482 | -0.3613 | 0.0277 | -0.0406 | 0.0298 | -0.1401 |
| MY | 0.0337 | 0.0475 | 0.0234 | 0.1215 | 0.0252 | 0.0391 |
| MY & MO | 0.0560 | -0.5807 | 0.0277 | -0.0402 | 0.0260 | 0.0054 |
| Model | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) |
| PR | 0.3802 | | 0.0432** | | 0.4121 | |
| FGT | 0.0320 | 0.0000*** | 0.0040** | 0.0018*** | 0.0547* | 0.0040** |
| MO | 0.7088 | 0.8324 | 0.0716* | 0.3136 | 0.5163 | 0.3780 |
| MY | 0.3252 | 0.0005*** | 0.0048** | 0.0014*** | 0.0551* | 0.0039*** |
| MY & MO | 0.4416 | 0.2981 | 0.0181** | 0.0094*** | 0.1369 | 0.0090*** |

**Panel B: 1947-2015: using an initial training period of $tp$ years.**

| training period ($tp$) | $tp$=20 | | $tp$=25 | | $tp$=30 | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | $R^2$ | OOS MSE | $R^2$ | OOS MSE | $R^2$ |
| HM | 0.0282 | 0 | 0.0293 | 0 | 0.0240 | 0 |
| PR | 0.0358 | -0.2717 | 0.0390 | -0.3281 | 0.0379 | -0.5788 |
| FGT | 0.0657 | -1.33270 | 0.0636 | -1.1691 | 0.0599 | -1.4965 |
| MO | 0.1591 | -4.6462 | 0.0470 | -0.6009 | 0.0405 | -0.6890 |
| MY | 0.0599 | -1.1266 | 0.0581 | -0.9790 | 0.0553 | -1.3047 |
| MY & MO | 0.1678 | -4.9548 | 0.0758 | -1.5836 | 0.0532 | -1.2168 |
| Model | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) |
| PR | 0.0535* | | 0.0973* | | 0.3030 | |
| FGT | 0.0666** | 0.0054*** | 0.1520 | 0.0051*** | 0.2808 | 0.0094*** |
| MO | 0.7970 | 0.0382** | 0.0.4716 | 0.0162** | 0.3502 | 0.0131** |
| MY | 0.0759* | 0.0083** | 0.1598 | 0.0069** | 0.2837 | 0.0127** |
| MY & MO | 0.8750 | 0.4698 | 0.7493 | 0.6488 | 0.4642 | 0.2374 |

$***$, $**$, $*$ significantly out-performs the benchmark forecasts at the 1% level, 5% level, and 10 % level, respectively. Reported p-values are one-sided. This table provides one-year out-of-sample forecasting results. The dependent variable in all cases are yearly log returns including dividends. Results for the time period 1901-2015 are shown in Panel A and results for 1947-2015 period are shown in Panel B. Column 2-3 reports the out-of-sample mean square error (OOS MSE) and out-of-sample $R^2$ (OOS $R^2$) with a 30/20-year training period for panel A/panel B. Column 4-5 shows OOS MSE and OOS $R^2$ with a 40/25-year training period for panel A/panel B. Column 6-7 gives OOS MSE and OOS $R^2$ with a 60/30-year training period for panel A/panel B. This table also provides the out-performance test results. HM is the out-of-sample historical mean. PR is the predictive regression model. FGT provides the estimates of (A1.4) in which the $\hat{x}_{t+1}$ is estimated by $x_t$ and $MY_t$. MO, MY, and MYMO provide estimates of (A1.4) using three different specifications of the demographic ratio projection $dr_{t+1|t}$ in (A1.3). The OOS $R^2$ refers to the out-of-sample $R^2$ using HM as the benchmark. CW is the Clark and West (2007) test. The columns marked (HM, CW) and (PR, CW) provide p-values for the CW test using HM and PR, respectively, as benchmarks.

Table B12: **Results of Return Regression Models (Pseudo Out-of-Sample, Rolling, Five-Year Ahead Forecast, IVX Estimation)**

**Panel A: 1901-2015: using a rolling window of length $w$**

| window ($w$) | $w = 30$ | | $w = 40$ | | $w = 60$ | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ |
| HM | 0.1698 | 0 | 0.1492 | 0 | 0.1266 | 0 |
| PR | 0.2028 | -0.1942 | 0.1775 | -0.1897 | 0.1273 | -0.0055 |
| FGT | 0.2594 | -0.5277 | 0.1988 | -0.3327 | 0.2641 | -1.0853 |
| MO | 0.4108 | -1.4194 | 0.1898 | -0.2721 | 0.2012 | -0.5886 |
| $MY_D$ | 0.2559 | -0.5071 | 0.1353 | 0.0930 | 0.1407 | -0.1107 |
| $MY_R$ | 0.1286 | 0.2428 | 0.1080 | 0.2757 | 0.0940 | 0.2578 |
| MY & MO | 0.6161 | -2.6287 | 0.3286 | -1.2028 | 0.3054 | -1.4111 |
| Model for $\widehat{dp}_{t+1\|t}$ | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) |
| PR | 0.8432 | | 0.8361 | | 0.5174 | |
| FGT | 0.8253 | 0.6974 | 0.7502 | 0.5893 | 0.9644 | 0.9793 |
| MO | 0.9836 | 0.9552 | 0.7137 | 0.5581 | 0.8392 | 0.8584 |
| $MY_D$ | 0.7175 | 0.6338 | 0.4128 | 0.2917 | 0.5910 | 0.5877 |
| $MY_R$ | 0.2571 | 0.1816 | 0.1964 | 0.1529 | 0.1842 | 0.1477 |
| MY & MO | 0.9635 | 0.9496 | 0.9646 | 0.9151 | 0.8941 | 0.9067 |

**Panel B: 1947-2015: using a rolling window of length $w$**

| window ($w$) | $w = 20$ | | $w = 25$ | | $w = 30$ | |
|---|---|---|---|---|---|---|
| Model for $\widehat{dp}_{t+1\|t}$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ |
| HM | 0.1941 | 0 | 0.1694 | 0 | 0.1465 | 0 |
| PR | 1.1298 | -4.8204 | 0.2727 | -0.6100 | 0.1742 | -0.1893 |
| FGT | 0.1599 | 0.1760 | 0.3184 | -0.8795 | 0.3079 | -1.1021 |
| MO | 1.0988 | -4.6605 | 0.4388 | -1.5904 | 0.3435 | -1.3454 |
| $MY_D$ | 0.1490 | 0.2324 | 0.0758 | 0.5527 | 0.0603 | 0.5880 |
| $MY_R$ | 0.0914 | 0.5291 | 0.0848 | 0.4991 | 0.0943 | 0.3561 |
| MY & MO | 3.7911 | -18.5305 | 0.8328 | -3.9159 | 0.4105 | -1.8024 |
| Model for $\widehat{dp}_{t+1\|t}$ | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) |
| PR | 0.9226 | | 0.9678 | | 0.9078 | |
| FGT | 0.2008 | 0.0728* | 0.8402 | 0.6184 | 0.8949 | 0.8395 |
| MO | 0.9592 | 0.4865 | 0.9979 | 0.9756 | 0.9780 | 0.9368 |
| $MY_D$ | 0.0691* | 0.0689* | 0.0029*** | 0.0029*** | 0.0040*** | 0.0007*** |
| $MY_R$ | 0.0075*** | 0.0607* | 0.0005*** | 0.0027*** | 0.0025*** | 0.0016*** |
| MY & MO | 0.9749 | 0.8994 | 0.9969 | 0.9890 | 0.9728 | 0.9429 |

$***$, $**$, $*$ significantly out-performs the benchmark forecasts at the 1% level, 5% level, and 10 % level, respectively. Reported p-values are one-sided. This table provides five-year out-of-sample forecasting results based on the rolling forecast method. It also provides the out-performance test results. The dependent variable in all cases are yearly log returns including dividends. Results for the time period 1901-2015 are shown in Panel A and results for 1947-2015 period are shown in Panel B. The dependent variable in all cases are yearly log returns including dividends. HM stands for historical mean. PR denotes the five-year ahead forecast obtained by forward recursion from the one-year ahead predictive regression forecast. $MY_D$ shows the results of (A1.8) by using MY as demographical choice with $k = 5$. FGT provides the estimates of (A1.4) in which the $\hat{x}_{t+1}$ is estimated by $x_t$ and $MY_t$. $MO$, $MY_R$ , and MYMO provide estimates of (A1.4) using three different specifications of the demographic ratio projection $dr_{t+h+1|t}$ in (A1.3) with $h = 4$. The OOS $R^2$ refers to the out-of-sample $R^2$ using HM as the benchmark. GW is the adjusted one-sided Giacomini and White (2006) test. The columns marked (HM, GW) and (PR, GW) provide p-values for the GW test using HM and PR, respectively, as benchmarks.

Table B13: **Results of Return Regression Models (Pseudo Out-of-Sample, Recursive, Five-Year Ahead Forecast, IVX Estimation)**

**Panel A: 1901-2015, forecasts begin $tp$ years after sample.**

| training period ($tp$) | $tp$=30 | | $tp$=40 | | $tp$=60 | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ |
| HM | 0.1438 | 0 | 0.1384 | 0 | 0.1171 | 0 |
| PR | 0.1339 | 0.0690 | 0.1349 | 0.0253 | 0.1205 | -0.0291 |
| FGT | 0.1125 | 0.2174 | 0.1155 | 0.1650 | 0.1150 | 0.0179 |
| MO | 0.2672 | -0.8588 | 0.1654 | -0.1952 | 0.1929 | -0.6475 |
| $MY_D$ | 0.1344 | 0.0654 | 0.1293 | 0.0654 | 0.0755 | 0.3555 |
| $MY_R$ | 0.1053 | 0.2677 | 0.0908 | 0.3442 | 0.0909 | 0.2233 |
| MY & MO | 0.2913 | -1.0260 | 0.1346 | 0.0271 | 0.1175 | -0.0033 |
| Model | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) |
| PR | 0.0764* | | 0.1221 | | 0.1808 | |
| FGT | 0.0202** | 0.0173** | 0.0411** | 0.0325** | 0.1148 | 0.1226 |
| MO | 0.5767 | 0.7727 | 0.1912 | 0.4533 | 0.4540 | 0.9266 |
| $MY_D$ | 0.0062*** | 0.0071*** | 0.0134** | 0.0149** | 0.0517* | 0.0665* |
| $MY_R$ | 0.0324** | 0.0299** | 0.0327** | 0.0222** | 0.0817* | 0.0797* |
| MY & MO | 0.4935 | 0.6605 | 0.0913* | 0.0914* | 0.1312 | 0.1250 |

**Panel B: 1947-2015: forecasts begin $tp$ years after sample.**

| training period ($tp$) | $tp$=20 | | $tp$=25 | | $tp$=30 | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | $R^2$ | OOS MSE | $R^2$ | OOS MSE | $R^2$ |
| HM | 0.1569 | 0 | 0.1286 | 0 | 0.1168 | 0 |
| PR | 0.1599 | -0.0196 | 0.1603 | -0.2465 | 0.1720 | -0.4727 |
| FGT | 0.2298 | -0.4652 | 0.2445 | -0.9011 | 0.2750 | -1.3548 |
| MO | 0.7401 | -3.7184 | 0.1618 | -0.2579 | 0.1532 | -0.3117 |
| $MY_D$ | 0.0986 | 0.3714 | 0.0866 | 0.3262 | 0.0903 | 0.2271 |
| $MY_R$ | 0.1344 | 0.1430 | 0.0992 | 0.2288 | 0.0997 | 0.1463 |
| MY & MO | 2.2532 | -13.3650 | 0.9047 | -6.0350 | 0.8311 | -6.1170 |
| Model | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) |
| PR | 0.1345 | | 0.2006 | | 0.2934 | |
| FGT | 0.0847* | 0.3325 | 0.1091 | 0.3773 | 0.1725 | 0.6799 |
| MO | 0.1258 | 0.0908* | 0.1203 | 0.1199 | 0.2271 | 0.1961 |
| $MY_D$ | 0.0568* | 0.0751* | 0.0564* | 0.0743* | 0.0853* | 0.1014 |
| $MY_R$ | 0.0495** | 0.1437 | 0.0387** | 0.1085 | 0.0792* | 0.1055 |
| MY & MO | 0.1558 | 0.3665 | 0.2881 | 0.6965 | 0.5900 | 0.8630 |

$***$, $**$, $*$ significantly out-performs the benchmark forecasts at the 1% level, 5% level, and 10 % level, respectively. Reported p-values are one-sided. This table provides one-year out-of-sample forecasting results. The dependent variable in all cases are yearly log returns including dividends. Results for the time period 1901-2015 are shown in Panel A and results for 1947-2015 period are shown in Panel B. Column 2-3 reports the out-of-sample mean square error (OOS MSE) and out-of-sample $R^2$ (OOS $R^2$) with a 30/20-year training period for panel A/panel B. Column 4-5 shows OOS MSE and OOS $R^2$ with a 40/25-year training period for panel A/panel B. Column 6-7 gives OOS MSE and OOS $R^2$ with a 60/30-year training period for panel A/panel B. This table also provides the out-performance test results. HM is the out-of-sample historical mean. PR is the predictive regression model. FGT provides the estimates of (A1.4) in which the $\hat{x}_{t+1}$ is estimated by $x_t$ and $MY_t$. MO, MY, and MYMO provide estimates of (A1.4) using three different specifications of the demographic ratio projection $dr_{t+1|t}$ in (A1.3). The OOS $R^2$ refers to the out-of-sample $R^2$ using HM as the benchmark. CW is the Clark and West (2007) test. The columns marked (HM, CW) and (PR, CW) provide p-values for the CW test using HM and PR, respectively, as benchmarks.

Table B14: **Results of Return Regression Models (Pseudo Out-of-Sample, Rolling, One-Year Ahead Forecast, IVX Estimation)**

**Panel A: 1901-2015: using a rolling window of length $w$.**

| window ($w$) | $w = 30$ | | $w = 40$ | | $w = 60$ | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ |
| HM | 0.0366 | 0 | 0.0272 | 0 | 0.0268 | 0 |
| PR | 0.0513 | -0.4030 | 0.0326 | -0.1992 | 0.0282 | -0.0511 |
| FGT | 0.0538 | -0.4706 | 0.0580 | -1.1354 | 0.0463 | -0.7265 |
| MO | 0.0636 | -0.7407 | 0.0439 | -0.6168 | 0.0617 | -1.3026 |
| MY | 0.0617 | -0.6863 | 0.0631 | -1.3256 | 0.0496 | -0.8492 |
| MY & MO | 0.0915 | -1.5024 | 0.0384 | -0.4142 | 0.0579 | -1.1588 |
| Model | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) |
| PR | 0.9850 | | 0.9372 | | 0.6493 | |
| FGT | 0.9654 | 0.5946 | 0.9702 | 0.9375 | 0.9694 | 0.9869 |
| MO | 0.9897 | 0.8701 | 0.9956 | 0.9398 | 0.9850 | 0.9873 |
| MY | 0.9988 | 0.8654 | 0.9883 | 0.9698 | 0.9776 | 0.9899 |
| MY & MO | 0.9900 | 0.9484 | 0.9750 | 0.8138 | 0.9783 | 0.9791 |

**Panel B: 1947-2015: using a rolling window of length $w$**

| window ($w$) | $w = 20$ | | $w = 25$ | | $w = 30$ | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ |
| HM | 0.0296 | 0 | 0.0308 | 0 | 0.0252 | 0 |
| PR | 0.0558 | -0.8840 | 0.0564 | -0.8303 | 0.0556 | -1.2031 |
| FGT | 0.1163 | -2.9293 | 0.0457 | -0.4838 | 0.0380 | -0.5055 |
| MO | 0.1658 | -4.5997 | 0.0438 | -0.4224 | 0.0342 | -0.3559 |
| MY | 0.0893 | -2.0153 | 0.0348 | -0.1281 | 0.0254 | -0.0079 |
| MY & MO | 0.3405 | -10.4998 | 0.1042 | -2.3805 | 0.0525 | -1.0798 |
| Model | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) |
| PR | 0.9481 | | 0.9348 | | 0.9937 | |
| FGT | 0.9999 | 0.9932 | 0.9221 | 0.2959 | 0.9593 | 0.1051 |
| MO | 0.9717 | 0.9346 | 0.8880 | 0.2597 | 0.8557 | 0.0640* |
| MY | 0.9998 | 0.9442 | 0.6798 | 0.1216 | 0.5194 | 0.0096*** |
| MY & MO | 0.9990 | 0.9971 | 0.9973 | 0.9339 | 0.9985 | 0.4176 |

$***$, $**$, $*$ significantly out-performs the benchmark forecasts at the 1% level, 5% level, and 10 % level, respectively. Reported p-values are one-sided. This table provides one-year out-of-sample forecasting results based on the rolling forecast method. The dependent variable in all cases are yearly log returns including dividends. Results for the time period 1901-2015 are shown in Panel A and results for 1947-2015 period are shown in Panel B. Column 2-3 reports the out-of-sample mean square error (OOS MSE) and out-of-sample $R^2$ (OOS $R^2$) with a 30/20-year training period for panel A/panel B. Column 4-5 shows OOS MSE and OOS $R^2$ with a 40/25-year training period for panel A/panel B. Column 6-7 gives OOS MSE and OOS $R^2$ with a 60/30-year training period for panel A/panel B. This table also provides the out-performance test results. HM is the out-of-sample historical mean. PR is the predictive regression model. FGT provides the estimates of (A1.4) in which the $\hat{x}_{t+1}$ is estimated by $x_t$and $MY_t$. MO, MY, and MYMO provide estimates of (A1.4) using three different specifications of the demographic ratio projection $dr_{t+1|t}$ in (A1.3). The OOS $R^2$ refers to the out-of-sample $R^2$ using HM as the benchmark. GW is the adjusted one-sided Giacomini and White (2006) test. The columns marked (HM, GW) and (PR, GW) provide p-values for the GW test using HM and PR, respectively, as benchmarks.

Table B15: **Results of Return Regression Models (True Out-of-Sample, Five-Year Ahead Unconditional Forecast, Real Time)**

**Panel A: Recursive, 1951-2015: using an initial training period of $tp$ years.**

| training period ($tp$) | $tp$=20 | | $tp$=25 | | $tp$=30 | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ |
| HM | 0.1270 | 0 | 0.1157 | 0 | 0.1273 | 0 |
| PR | 0.1531 | -0.2048 | 0.1629 | -0.4074 | 0.1684 | -0.3233 |
| FGT | 0.1784 | -0.4043 | 0.1871 | -0.6169 | 0.2036 | -0.6000 |
| MO | 0.1703 | -0.3408 | 0.1793 | -0.5496 | 0.2028 | -0.5934 |
| $MY_D$ | **0.0989** | 0.2219 | **0.1082** | 0.0650 | **0.1213** | 0.0468 |
| $MY_R$ | 0.1059 | 0.1662 | 0.1144 | 0.0113 | 0.1283 | -0.0078 |
| MY & MO | 0.2350 | -0.8495 | 0.2378 | -1.0551 | 0.2701 | -1.1222 |
| Model | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) |
| PR | 0.1996 | | 0.3341 | | 0.3110 | |
| FGT | 0.1395 | 0.1341 | 0.2681 | 0.1931 | 0.3625 | 0.4877 |
| MO | 0.1048 | 0.1738 | 0.2049 | 0.2571 | 0.2354 | 0.3436 |
| $MY_D$ | 0.0561* | 0.0804* | 0.0880* | 0.1117 | 0.1174 | 0.1700 |
| $MY_R$ | 0.0863* | 0.0558* | 0.1560 | 0.0817* | 0.1914 | 0.1465 |
| MY & MO | 0.2787 | 0.4018 | 0.2855 | 0.3835 | 0.2938 | 0.4537 |

**Panel B: Rolling, 1951-2015: using an initial training period of $tp$ years.**

| training period ($tp$) | $tp$=20 | | $tp$=25 | | $tp$=30 | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | $R^2$ | OOS MSE | $R^2$ | OOS MSE | $R^2$ |
| HM | 0.1754 | 0 | 0.1593 | 0 | 0.1550 | 0 |
| PR | 0.2414 | -0.3761 | 0.2697 | -0.6934 | 0.2137 | -0.3786 |
| FGT | 0.2309 | -0.3164 | 0.3228 | -1.0265 | 0.2653 | -0.7116 |
| MO | 0.1889 | -0.0769 | 0.1530 | 0.0396 | 0.1904 | -0.2278 |
| $MY_D$ | 0.1784 | -0.0170 | **0.0971** | 0.3907 | **0.0858** | 0.4468 |
| $MY_R$ | **0.0971** | 0.4465 | 0.1076 | 0.3245 | 0.1223 | 0.2113 |
| MY & MO | 0.1926 | -0.0977 | 0.1414 | 0.1120 | 0.1675 | -0.0805 |
| Model for $\widehat{dp}_{t+1\|t}$ | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) |
| PR | 0.7981 | | 0.9099 | | 0.8639 | |
| FGT | 0.9163 | 0.4489 | 0.9822 | 0.8576 | 0.9807 | 0.8920 |
| MO | 0.5567 | 0.3547 | 0.4697 | 0.2033 | 0.6258 | 0.4337 |
| $MY_D$ | 0.5347 | 0.2512 | 0.0116** | 0.0464** | 0.0313** | 0.0668* |
| $MY_R$ | 0.0009*** | 0.0508* | 0.0093*** | 0.0454** | 0.0760* | 0.0713* |
| MY & MO | 0.5823 | 0.3544 | 0.3974 | 0.1641 | 0.5539 | 0.3578 |

$***$, $**$, $*$ significantly out-performs the benchmark forecasts at the 1% level, 5% level, and 10 % level, respectively. Reported p-values are one-sided. This table provides five-year out-of-sample unconditional forecasting results using the real time forecasts. The dependent variable in all cases are yearly log returns including dividends. Results of recursive method are shown in Panel A and results of rolling method are shown in Panel B. Column 2-3 reports the out-of-sample mean square error (OOS MSE) and out-of-sample $R^2$ (OOS $R^2$) with a 20-year training period for panel A/panel B. Column 4-5 shows OOS MSE and OOS $R^2$ with a 25-year training period for panel A/panel B. Column 6-7 gives OOS MSE and OOS $R^2$ with a 30-year training period for panel A/panel B. This table also provides the out-performance test results. HM is the out-of-sample historical mean. PR is the predictive regression model. FGT provides the estimates of (A1.4) in which the $\hat{x}_{t+1}$ is estimated by $x_t$ and $MY_t$. MO, MY, and MYMO provide estimates of (A1.4) using three different specifications of the demographic ratio projection $dr_{t+1|t}$ in (A1.3). The OOS $R^2$ refers to the out-of-sample $R^2$ using HM as the benchmark CW is the Clark and West (2007) test. GW is the adjusted one-sided Giacomini and White (2006) test. The columns marked (HM, CW) and (PR, CW) provide p-values for the CW test using HM and PR, respectively, as benchmarks. Similarly columns mark (HM,GW) and (PR,GW) provide p-values for the GW test.

Table B16: **Results of Return Regression Models (True Out-of-Sample, One-Year Ahead Forecast, Stale)**

**Panel A: Recursive, 1951-2015: using an initial training period of $tp$ years.**

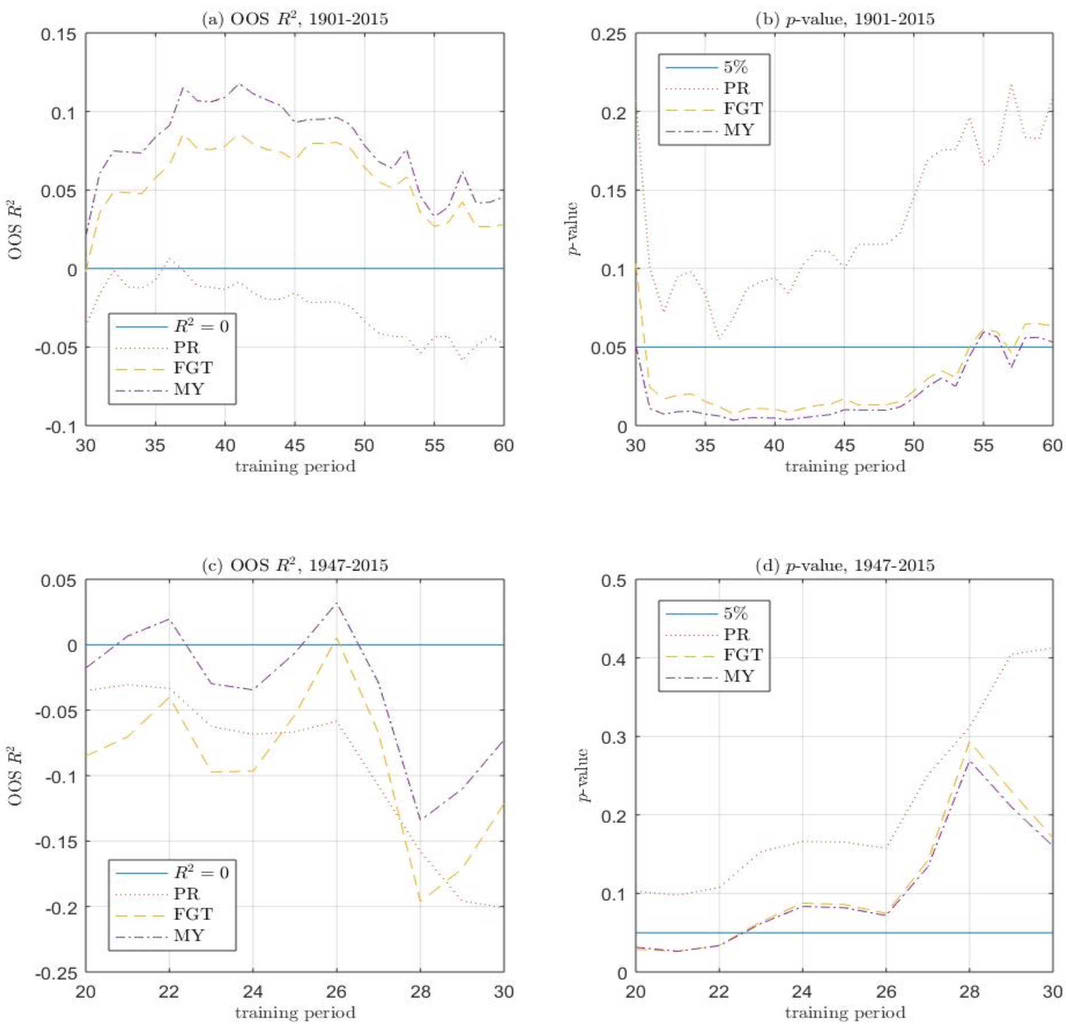
| training period ($tp$) | $tp$=20 | | $tp$=25 | | $tp$=30 | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ | OOS MSE | OOS $R^2$ |
| HM | 0.0284 | 0 | 0.0237 | 0 | 0.0253 | 0 |
| PR | 0.0313 | -0.1003 | 0.0292 | -0.2312 | 0.0302 | -0.1908 |
| FGT | 0.0318 | -0.1190 | 0.0283 | -0.1936 | 0.0305 | -0.2020 |
| MO | 0.0345 | -0.2120 | 0.0304 | -0.2821 | 0.0329 | -0.2967 |
| MY | 0.0297 | -0.0430 | 0.0265 | -0.1187 | 0.0285 | -0.1246 |
| MY & MO | 0.0340 | -0.1964 | 0.0289 | -0.2197 | 0.0308 | -0.2145 |
| Model | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) | $P$ Value (HM, CW) | $P$ Value (PR, CW) |
| PR | 0.2105 | | 0.5170 | | 0.5312 | |
| FGT | 0.1498 | 0.0520* | 0.3670 | 0.0379** | 0.4062 | 0.1612 |
| MO | 0.4011 | 0.1596 | 0.4237 | 0.1167 | 0.4306 | 0.2569 |
| MY | 0.1053 | 0.0211** | 0.2717 | 0.0135** | 0.2973 | 0.0585* |
| MY & MO | 0.5516 | 0.4180 | 0.4483 | 0.1216 | 0.4395 | 0.2362 |

**Panel B: Rolling, 1951-2015, forecasts begin $tp$ years after sample.**

| training period ($tp$) | $tp$=20 | | $tp$=25 | | $tp$=30 | |
|---|---|---|---|---|---|---|
| Model | OOS MSE | $R^2$ | OOS MSE | $R^2$ | OOS MSE | $R^2$ |
| HM | 0.0303 | 0 | 0.0252 | 0 | 0.0265 | 0 |
| PR | 0.0302 | 0.0030 | 0.0312 | -0.2405 | 0.0300 | -0.1318 |
| FGT | 0.0289 | 0.0461 | 0.0257 | -0.0221 | 0.0289 | -0.0895 |
| MO | 0.0315 | -0.0419 | 0.0283 | -0.1239 | 0.0309 | -0.1631 |
| MY | 0.0286 | 0.0546 | 0.0250 | 0.0065 | 0.0249 | 0.0635 |
| MY & MO | 0.0379 | -0.2532 | 0.0327 | -0.2977 | 0.0339 | -0.2791 |
| Model for $\widehat{dp}_{t+1\|t}$ | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) | $P$ Value (HM, GW) | $P$ Value (PR, GW) |
| PR | 0.4897 | | 0.9704 | | 0.9325 | |
| FGT | 0.3486 | 0.3670 | 0.5854 | 0.1050 | 0.7876 | 0.3916 |
| MO | 0.5812 | 0.5824 | 0.7121 | 0.3222 | 0.7376 | 0.5464 |
| MY | 0.3148 | 0.3335 | 0.4763 | 0.0700* | 0.2551 | 0.0862* |
| MY & MO | 0.9202 | 0.8924 | 0.9408 | 0.5958 | 0.9369 | 0.7583 |

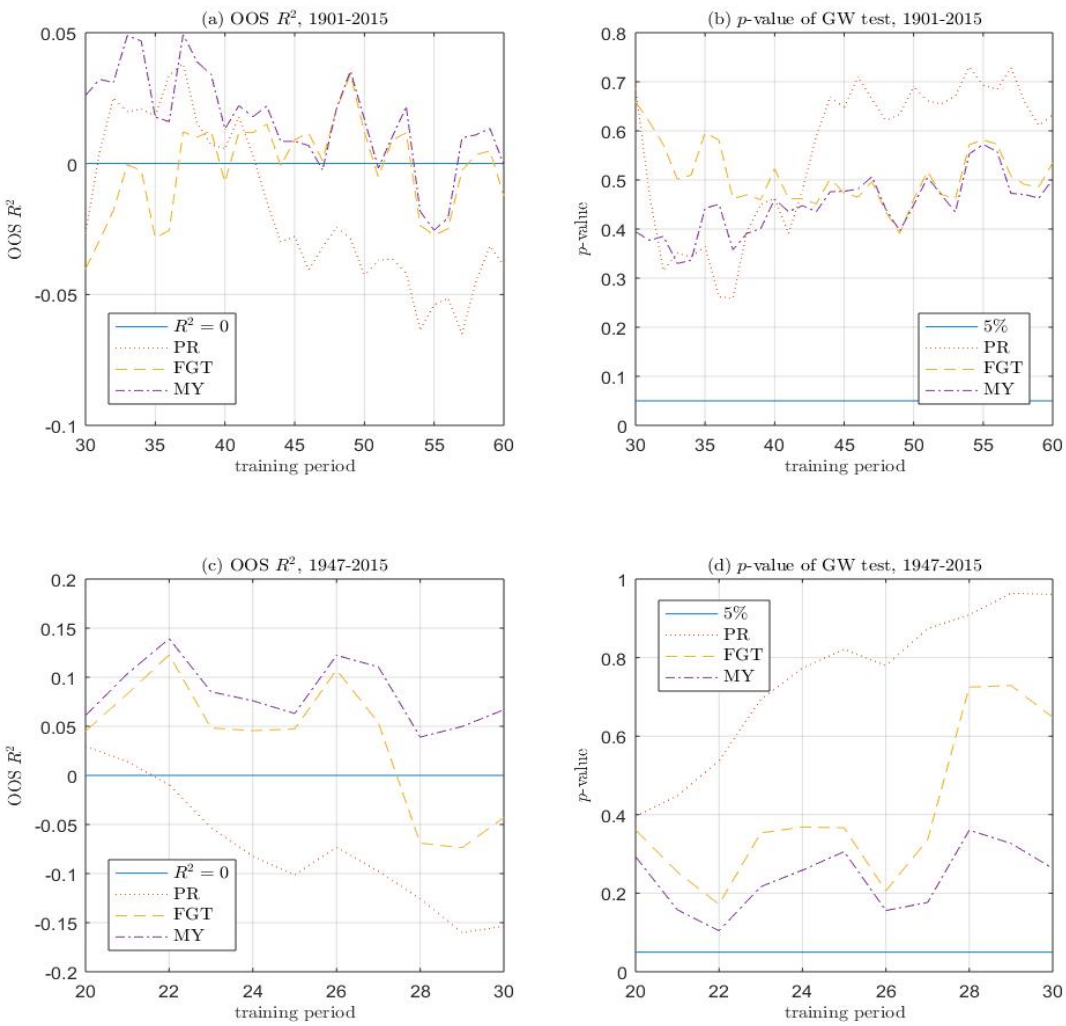
$***$, $**$, $*$ significantly out-performs the benchmark forecasts at the 1% level, 5% level, and 10 % level, respectively. Reported p-values are one-sided. This table provides one-year out-of-sample unconditional forecasting results using the stale forecasts. The dependent variable in all cases are yearly log returns including dividends. Results of recursive method are shown in Panel A and results of rolling method are shown in Panel B. Column 2-3 reports the out-of-sample mean square error (OOS MSE) and out-of-sample $R^2$ (OOS $R^2$) with a 20-year training period for panel A/panel B. Column 4-5 shows OOS MSE and OOS $R^2$ with 25-year training period for panel A/panel B. Column 6-7 gives OOS MSE and OOS $R^2$ with a 30-year training period for panel A/panel B. This table also provides the out-performance test results. HM is the out-of-sample historical mean. PR is the predictive regression model. FGT provides the estimates of (A1.4) in which the $\hat{x}_{t+1}$ is estimated by $x_t$ and $MY_t$. MO, MY, and MYMO provide estimates of (A1.4) using three different specifications of the demographic ratio projection $dr_{t+1|t}$ in (A1.3). The OOS $R^2$ refers to the out-of-sample $R^2$ using HM as the benchmark. CW is the Clark and West (2007) test. GW is the adjusted one-sided Giacomini and White (2006) test. The columns marked (HM, CW) and (PR, CW) provide p-values for the CW test using HM and PR, respectively, as benchmarks. Similarly columns mark (HM,GW) and (PR,GW) provide p-values for the GW test.

Figure B1: Conditional Pseudo Out-of-Sample Forecasts, One-Year Ahead



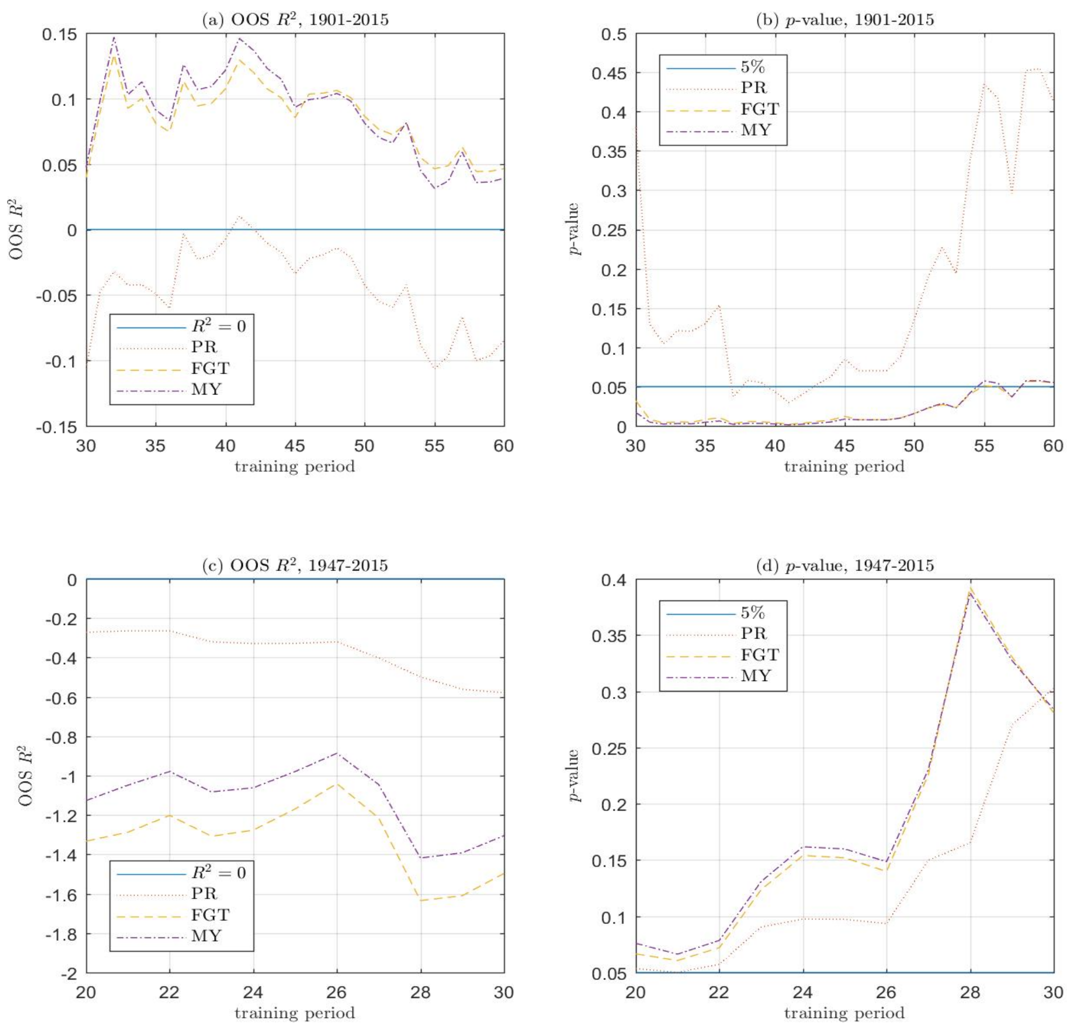

This figure provides one-year ahead out-of-sample forecasting results with respect to different window sizes based on the recursive method. (a)-(b) show the OOS $R^2$s and one-sided p-values for CW test using the full sample period 1901-2015 with the window size varying from 30-60 years. (c)-(d) show the analogous result for the post WWII period 1947-2015 with the window size varying from 20-30 years. The purple, red, and yellow dashed lines stand for the $MY$, predictive regression model, and FGT model respectively. The blue solid line stands for the historical mean model in (a) and (c) and the 5% significant level in (b) and (d). The OOS $R^2$ uses the HM as its benchmark and is defined in Footnote 11.

Figure B2: Pseudo Out-of-Sample Rolling Forecasts, One-Year Ahead



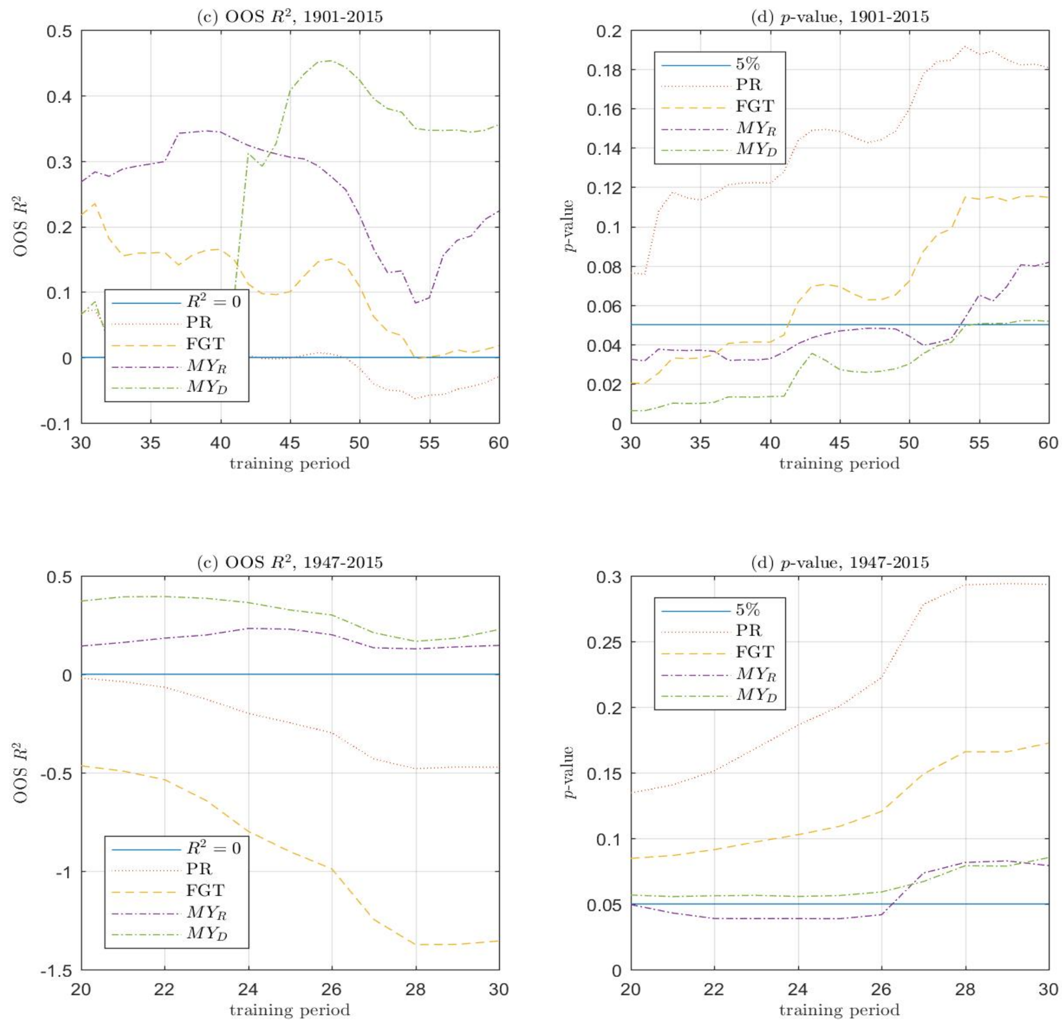

This figure provides one-year ahead out-of-sample forecasting results with respect to different window size based on the rolling method. (a)-(b) show the OOS $R^2$s and p-values for GW test for the full sample period 1901-2015 with the window size varying from 30-60 years. (c)-(d) show the analogous result for the post WWII period 1947-2015 with the window size varying from 20-30 years. The purple, red, and yellow dashed lines stand for the $MY$, predictive regression model, and FGT model respectively. The blue solid line stands for the historical mean model in (a) and (c) and the 5% significant level in (b) and (d). The OOS $R^2$ uses the HM as its benchmark and is defined in Footnote 11.

Figure B3: Pseudo Out-of-Sample Recursive Forecasts, One-Year Ahead, IVX Estimation



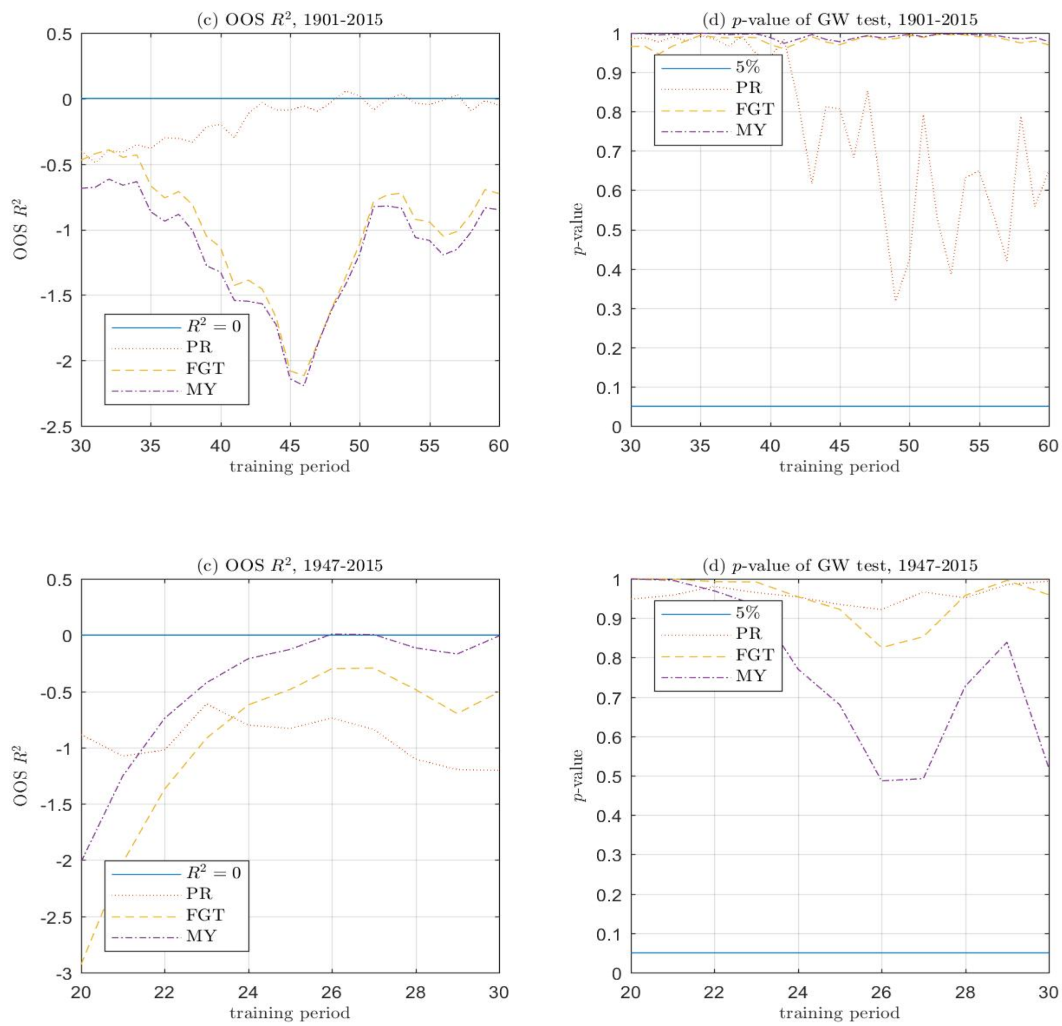

This figure provides IVX estimates of one-year ahead out-of-sample forecasting results with respect to different window sizes based on the recursive method. (a)-(b) show the OOS $R^2$s and one-sided p-values for the CW test for the full sample period 1901-2015 with the window size varying from 30-60 years. (c)-(d) show the analogous result for the post WWII period 1947-2015 with the window size varying from 20-30 years. The purple, red, and yellow dashed lines stand for the $MY$, predictive regression model, and FGT model respectively. The blue solid line stands for the historical mean model in (a) and (c) and the 5% significant level in (b) and (d). The OOS $R^2$ uses the HM as its benchmark and is defined in Footnote 11.

Figure B4: Pseudo Out-of-Sample Recursive Forecasts, Five-Year Ahead, IVX Estimation



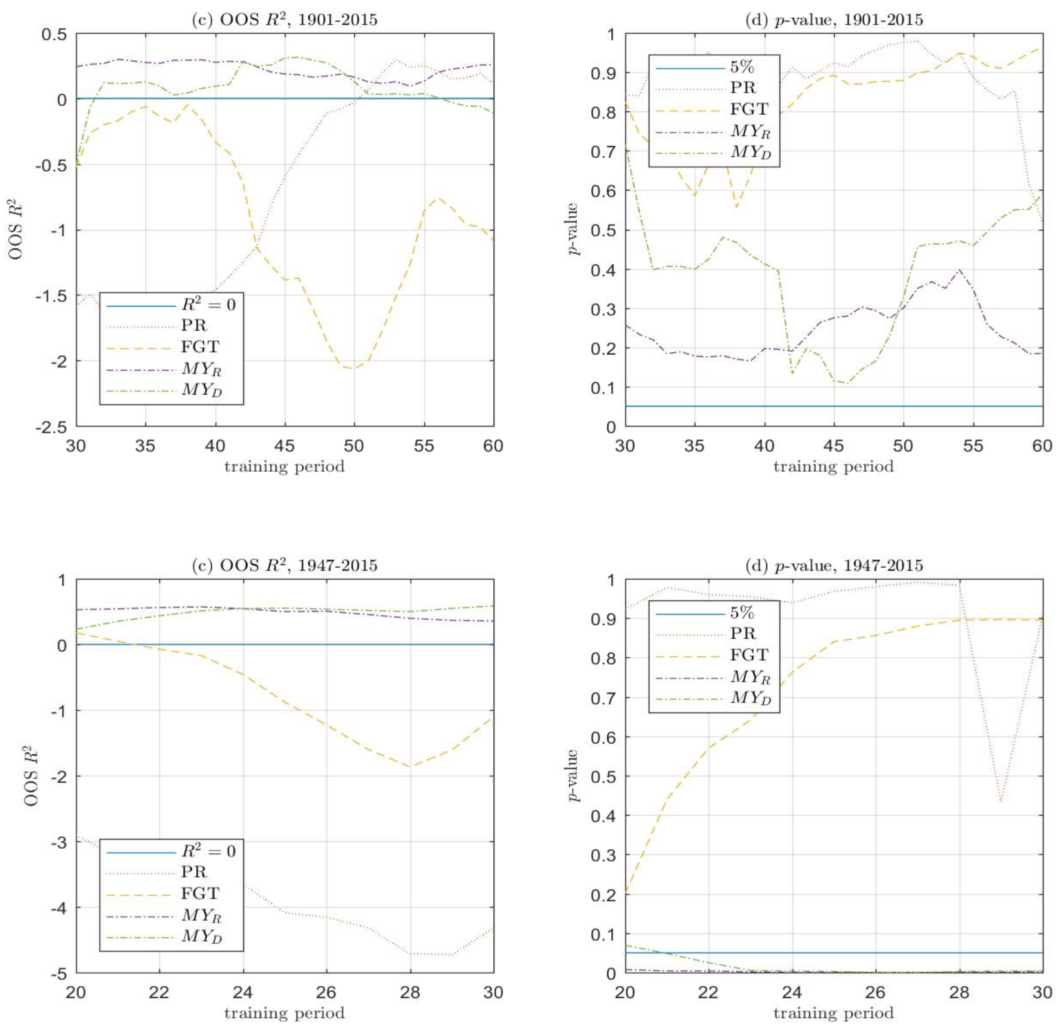

This figure provides IVX estimates of five-year ahead out-of-sample forecasting results with respect to different window size based on the recursive method. (a)-(b) show the OOS $R^2$s and one-sided p-values for the CW test for the full sample period 1901-2015 with the window size varying from 30-60 years. (c)-(d) show the analogous result for the post WWII period 1947-2015 with the window size varying from 20-30 years. The green, purple, red, and yellow dashed lines stand for the $MY_D$, $MY_R$, predictive regression model, and FGT model respectively. The blue solid line stands for the historical mean model in (a) and (c) and the 5% significant level in (b) and (d). The OOS $R^2$ uses the HM as its benchmark and is defined in Footnote 11.

Figure B5: Pseudo Out-of-Sample Rolling Forecasts, One-Year Ahead, IVX Estimation



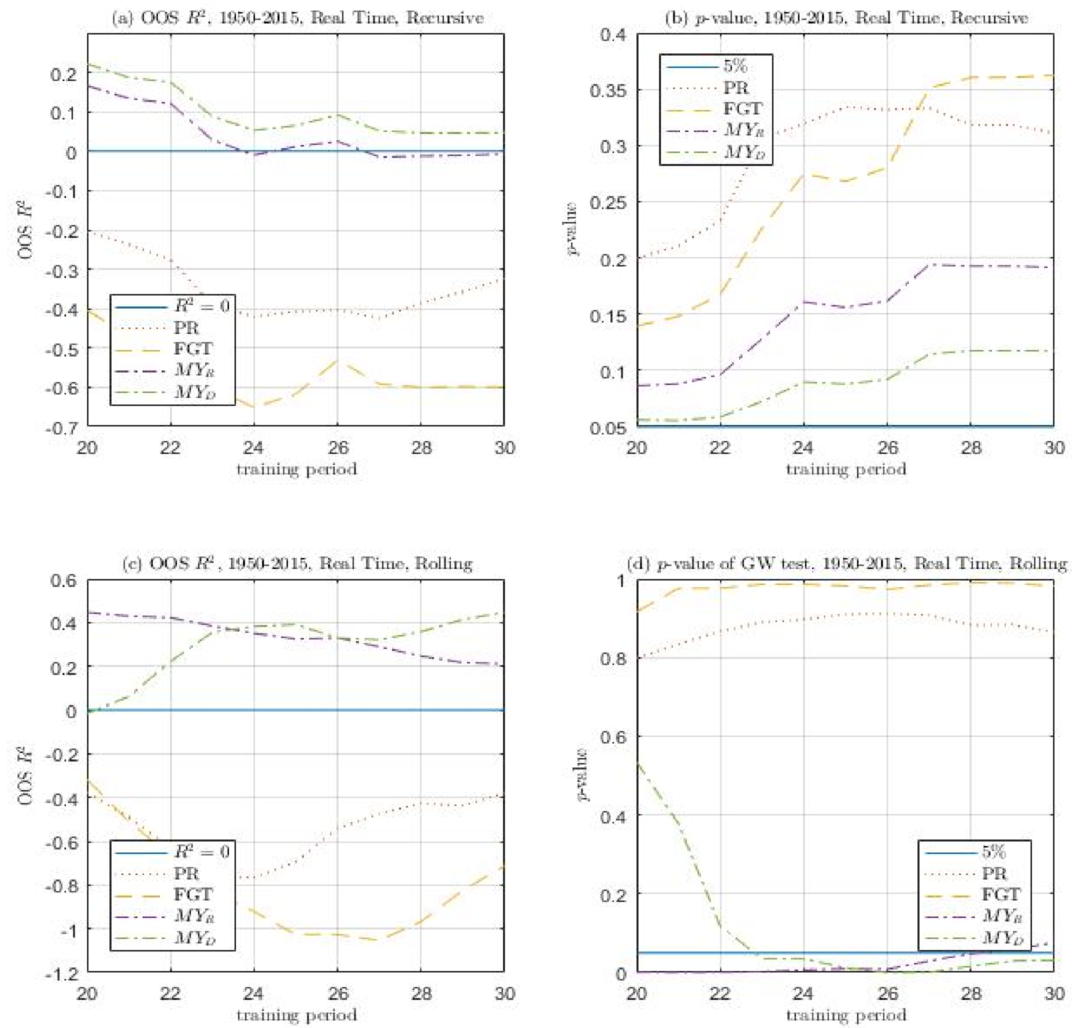

This figure provides IVX estimates of one-year ahead out-of-sample forecasting results with respect to different window size based on the rolling method. (a)-(b) show the OOS $R^2$s and p-values for the GW test for the full sample period 1901-2015 with the window size varying from 30-60 years. (c)-(d) show the analogous result for the post WWII period 1947-2015 with the window size varying from 20-30 years. The purple, red, and yellow dashed lines stand for the $MY$, predictive regression model, and FGT model respectively. The blue solid line stands for the historical mean model in (a) and (c) and the 5% significant level in (b) and (d). The OOS $R^2$ uses the HM as its benchmark and is defined in Footnote 11.

Figure B6: Pseudo Out-of-Sample Rolling Forecasts, Five-Year Ahead, IVX Estimation



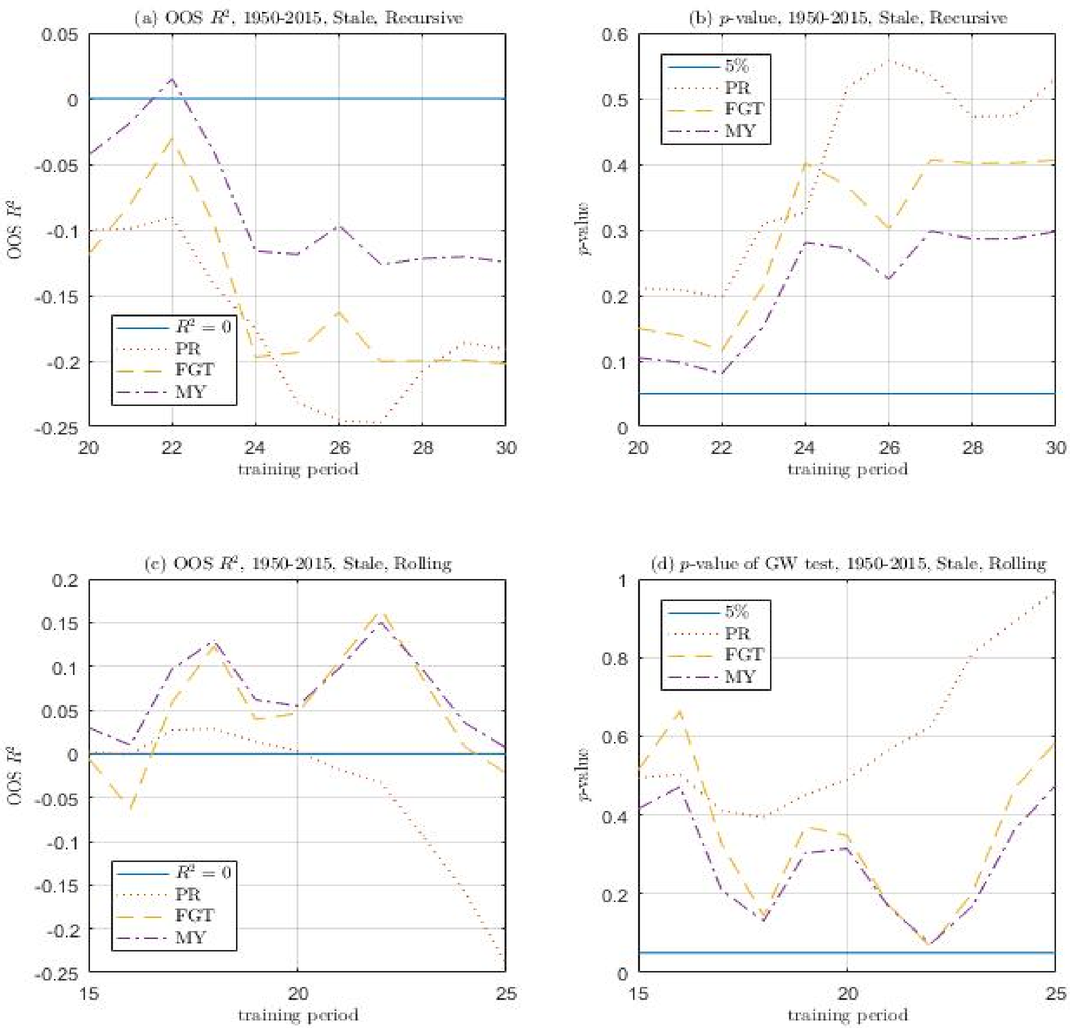

This figure provides IVX estimates of five-year ahead out-of-sample forecasting results with respect to different window size based on the rolling method. (a)-(b) show the OOS $R^2$s and p-values for the GW test for the full sample period 1901-2015 with the window size varying from 30-60 years. (c)-(d) show the analogous result for the post WWII period 1947-2015 with the window size varying from 20-30 years. The green, purple, red, and yellow dashed lines stand for the $MY_D$, $MY_R$, predictive regression model, and FGT model respectively. The blue solid line stands for the historical mean model in (a) and (c) and the 5% significant level in (b) and (d). The OOS $R^2$ uses the HM as its benchmark and is defined in Footnote 11.

Figure B7: True Out-of-Sample Forecasts, Five-Year Ahead, Real Time



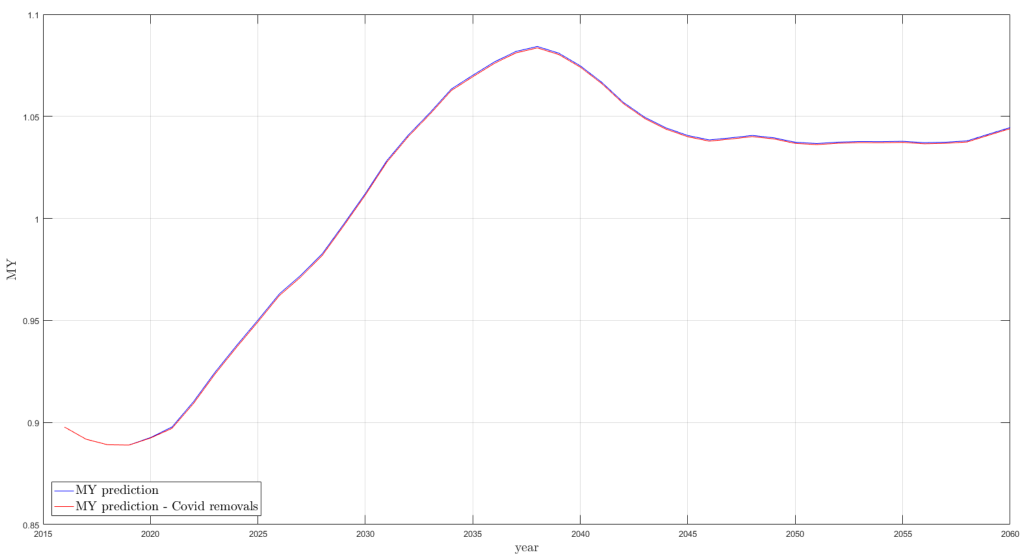

This figure provides five-year ahead out-of-sample forecasting results with respect to different window size using real time forecasts. (a)-(b) show the OOS $R^2$s and p-values for the CW test with the recursive method. (c)-(d) show the analogous results for the GW test with the rolling method. The green, purple, red, and yellow dashed lines stand for the $MY_D$, $MY_R$, predictive regression model, and FGT model respectively. The blue solid line stands for the historical mean model in (a) and (c) and the 5% significant level in (b) and (d). The OOS $R^2$ uses the HM as its benchmark and is defined in Footnote 11.

Figure B8: True Out-of-Sample Forecasts, One-Year Ahead, Stale



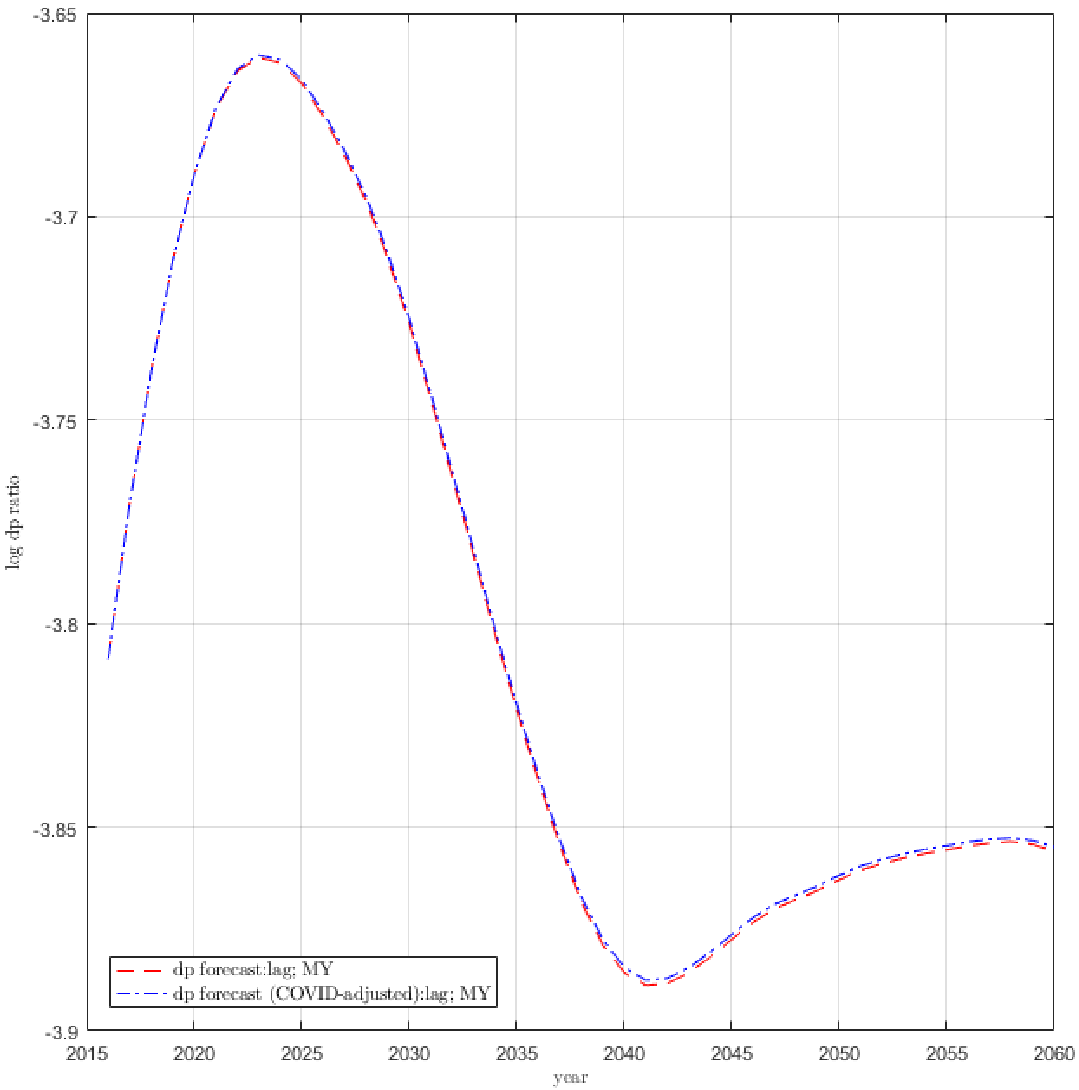

This figure provides one-year ahead out-of-sample forecasting results with respect to different window sizes using stale forecasts. (a)-(b) show the OOS $R^2$s and one-sided p-values for the CW test with the recursive method. (c)-(d) show the analogous results for the GW test with the rolling method. The purple, red, and yellow dashed lines stand for the $MY$, predictive regression model, and FGT model respectively. The blue solid line stands for the historical mean model in (a) and (c) and the 5% significant level in (b) and (d). The OOS $R^2$ uses the HM as its benchmark and is defined in Footnote 11.

Figure B9: Original and Covid-19 Adjusted MY Ratio Annual Projection, Scenario One: $\theta = 10\%, t^* = 2023$



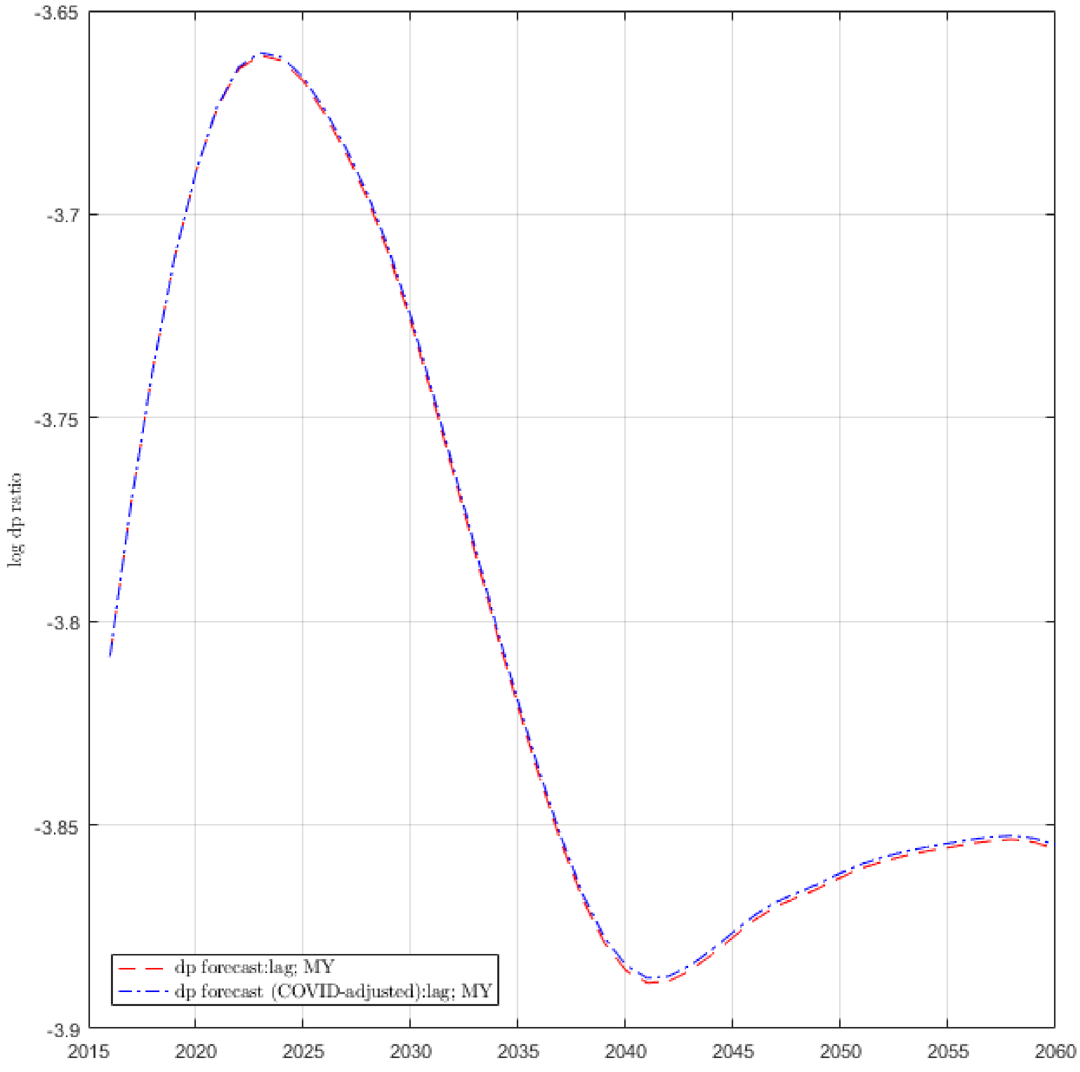

This figure provides MY ratio projections for both unadjusted and Covid-19 adjusted annual projections. The red line presents the MY ratio of the projection from Census Bureau 2017. The blue line shows the MY ratio Covid-19 adjusted projection.

Figure B10: Projections for the dp Ratio Based on Covid19 Adjusted Demographic Ratios, Scenario One: $\theta = 10\%, t^* = 2023$



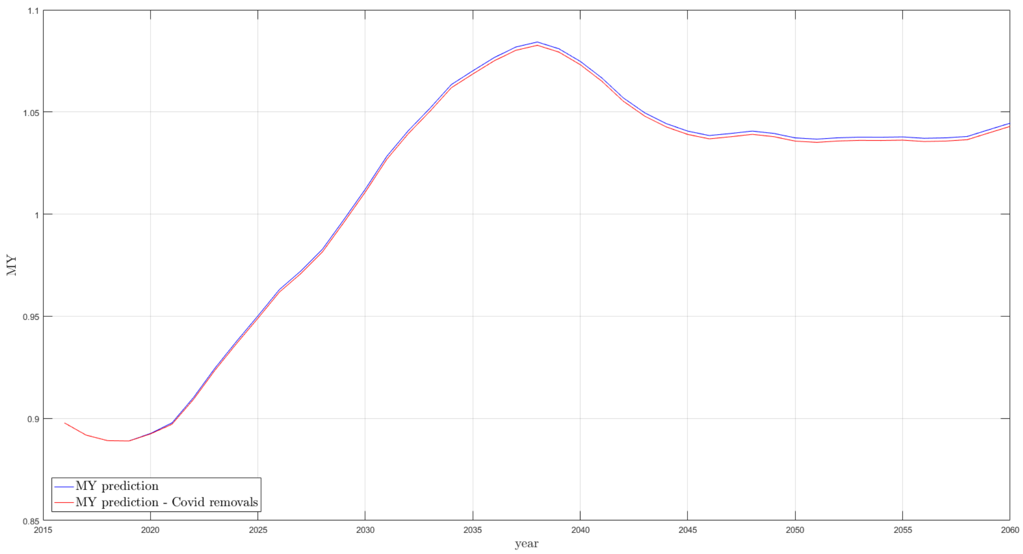

This figure provides forecasts for the dividend-price ratio based on both the Census Bureau projection and the Covid-19 adjusted projection. The red dashed line shows the projection based on the Census Bureau projection. The blue dashed line gives the dp ratio projection from the Covid-19 adjusted demographic projection.

Figure B11: Covid19 Adjusted Projected Five-Year Rolling Average Stock Returns, Scenario One: $\theta = 10\%, t^* = 2023$



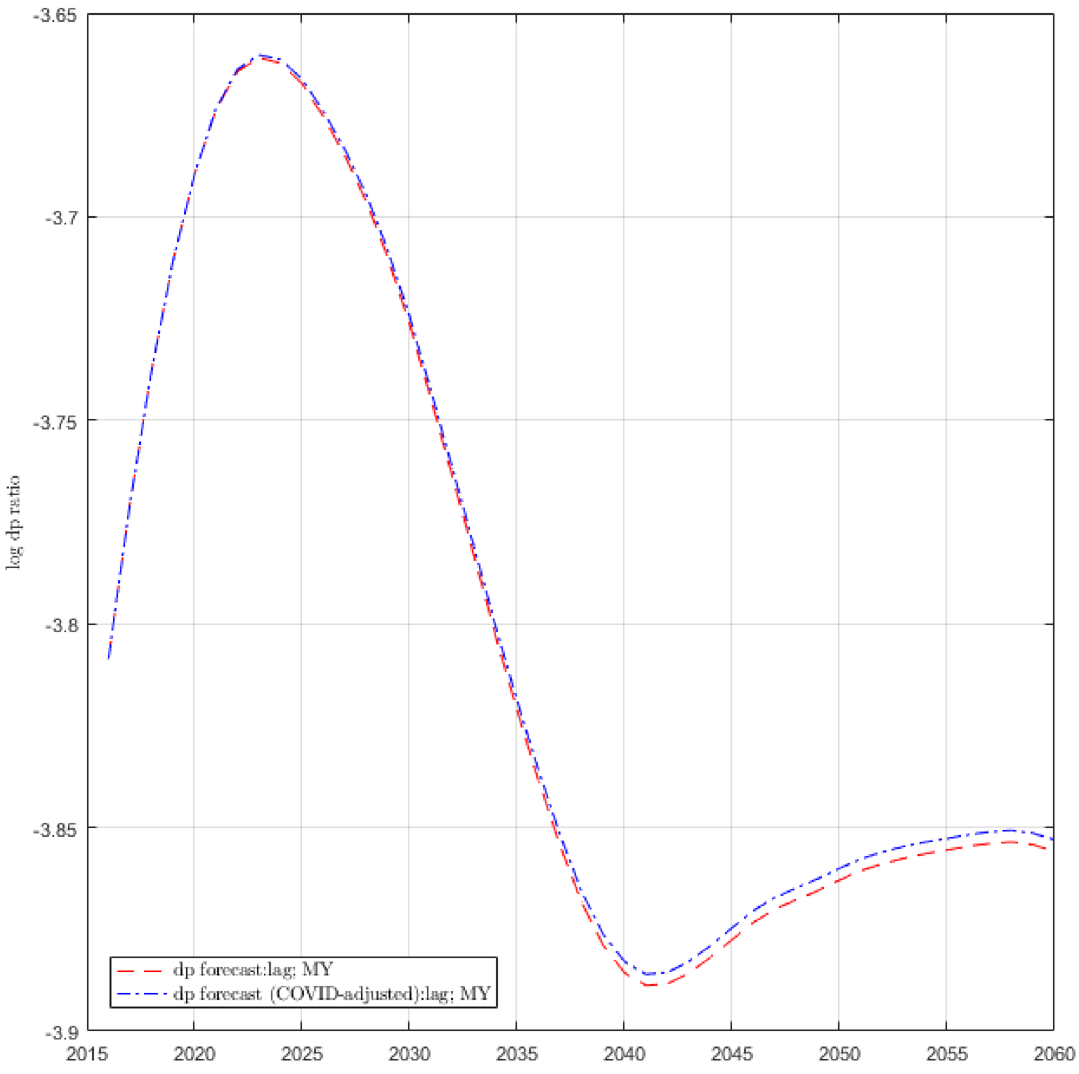

This figure provides forecasts for the five year rolling average return (including dividends) on the S&P 500 index using both the Census Bureau demographic projections and the Covid-19 adjusted demographic projection. The red dashed line shows the projection based on the Census Bureau projection. The blue dashed line gives the projection based on the Covid-19 adjusted demographic projection.

Figure B12: Original and Covid-19 Adjusted MY Ratio Annual Projection, Scenario Two: $\theta = 30\%, t^* = 2023$



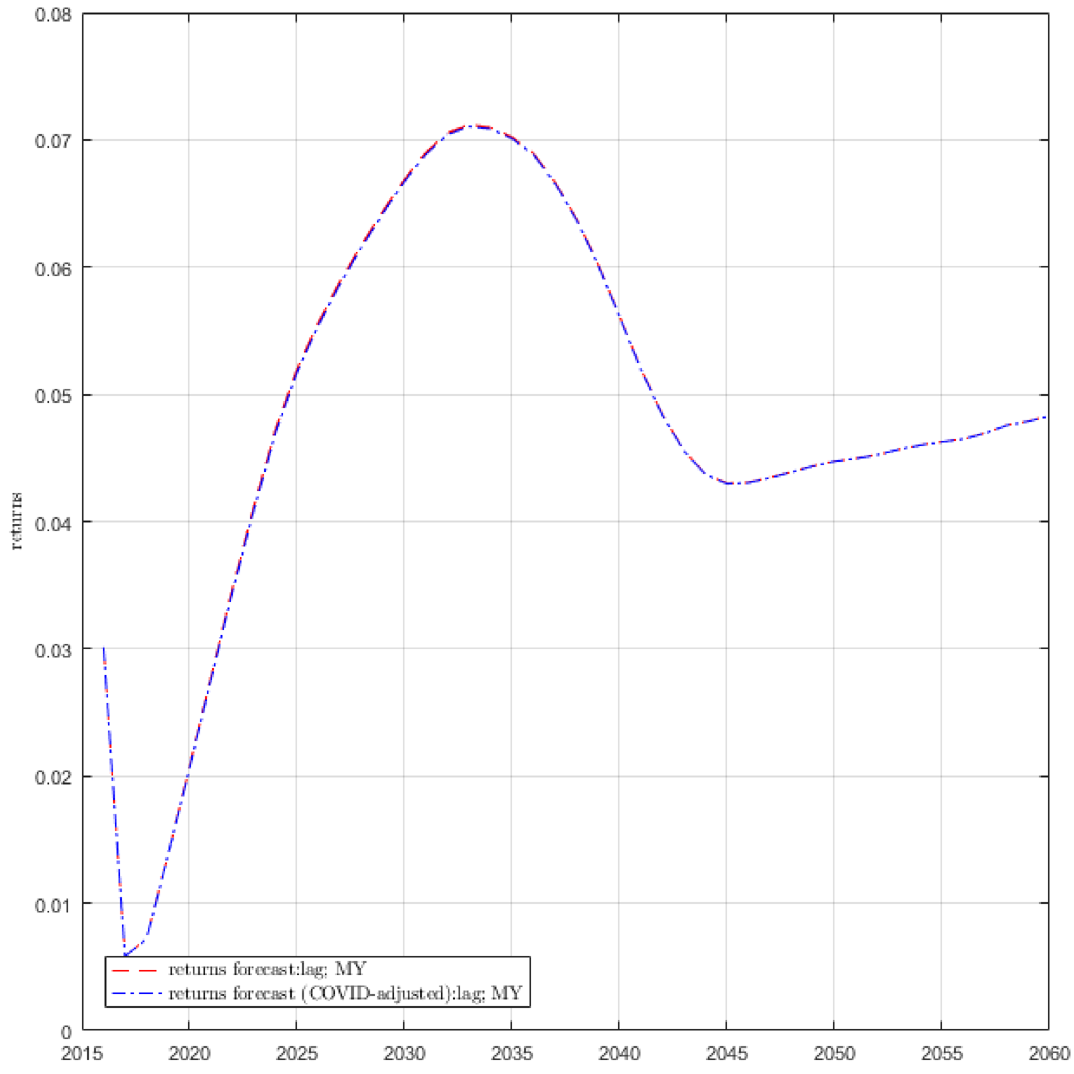

This figure provides MY ratio projections for both unadjusted and Covid-19 adjusted annual projections. The red line presents the MY ratio of the projection from Census Bureau 2017. The blue line shows the MY ratio Covid-19 adjusted projection.

Figure B13: Projections for the dp Ratio Based on Covid19 Adjusted Demographic Ratios, Scenario Two: $\theta = 30\%, t^* = 2023$



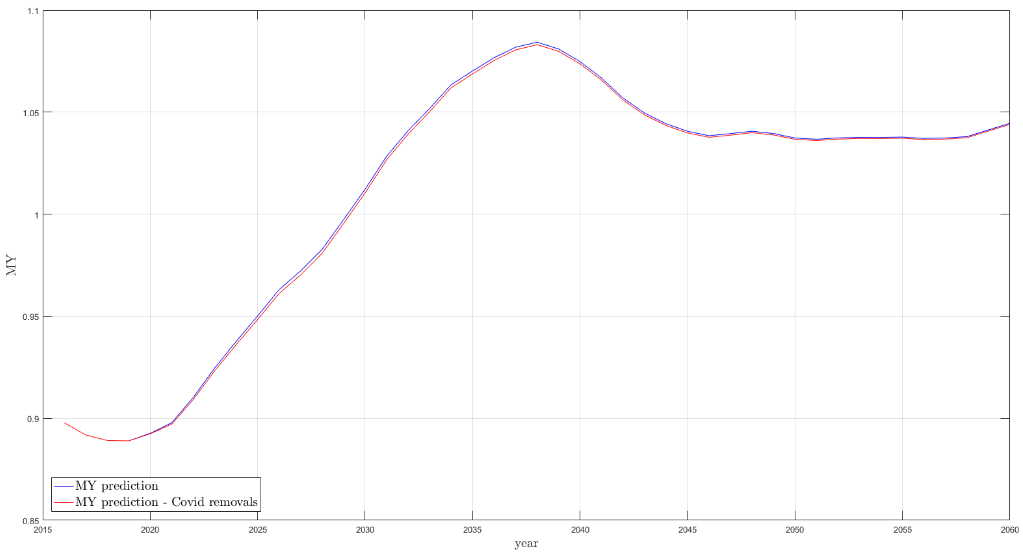

This figure provides forecasts for the dividend-price ratio based on both the Census Bureau projection and the Covid-19 adjusted projection. The red dashed line shows the projection based on the Census Bureau projection. The blue dashed line gives the dp ratio projection from the Covid-19 adjusted demographic projection.

Figure B14: Covid19 Adjusted Projected Five-Year Rolling Average Stock Returns, Scenario Two: $\theta = 30\% t^* = 2023$



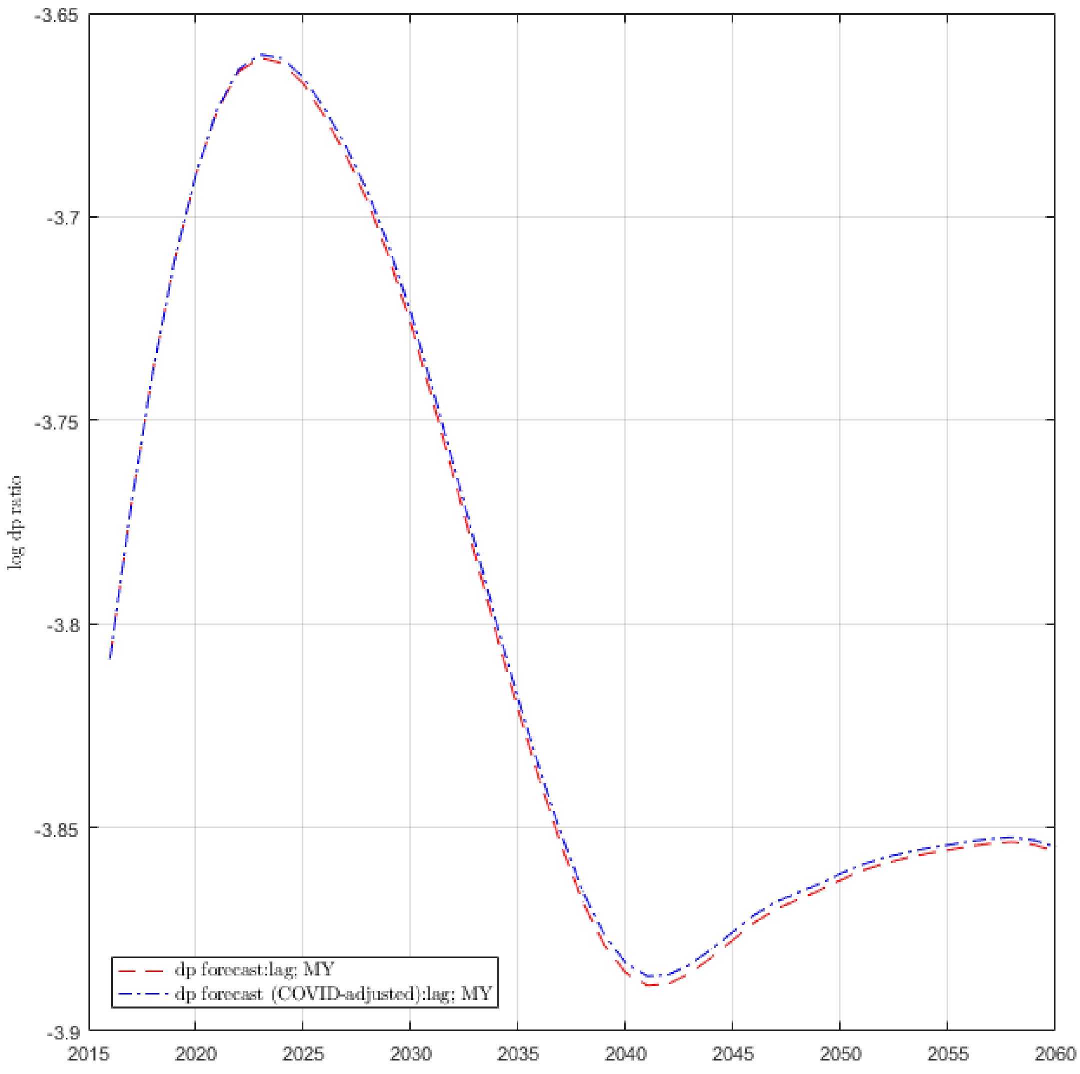

This figure provides forecasts for the five year rolling average return (including dividends) on the S&P 500 index using both the Census Bureau demographic projections and the Covid-19 adjusted demographic projection. The red dashed line shows the projection based on the Census Bureau projection. The blue dashed line gives the projection based on the Covid-19 adjusted demographic projection.

Figure B15: Original and Covid-19 Adjusted MY Ratio Annual Projection, Scenario Three: $\theta = 10\%, t^* = 2030$



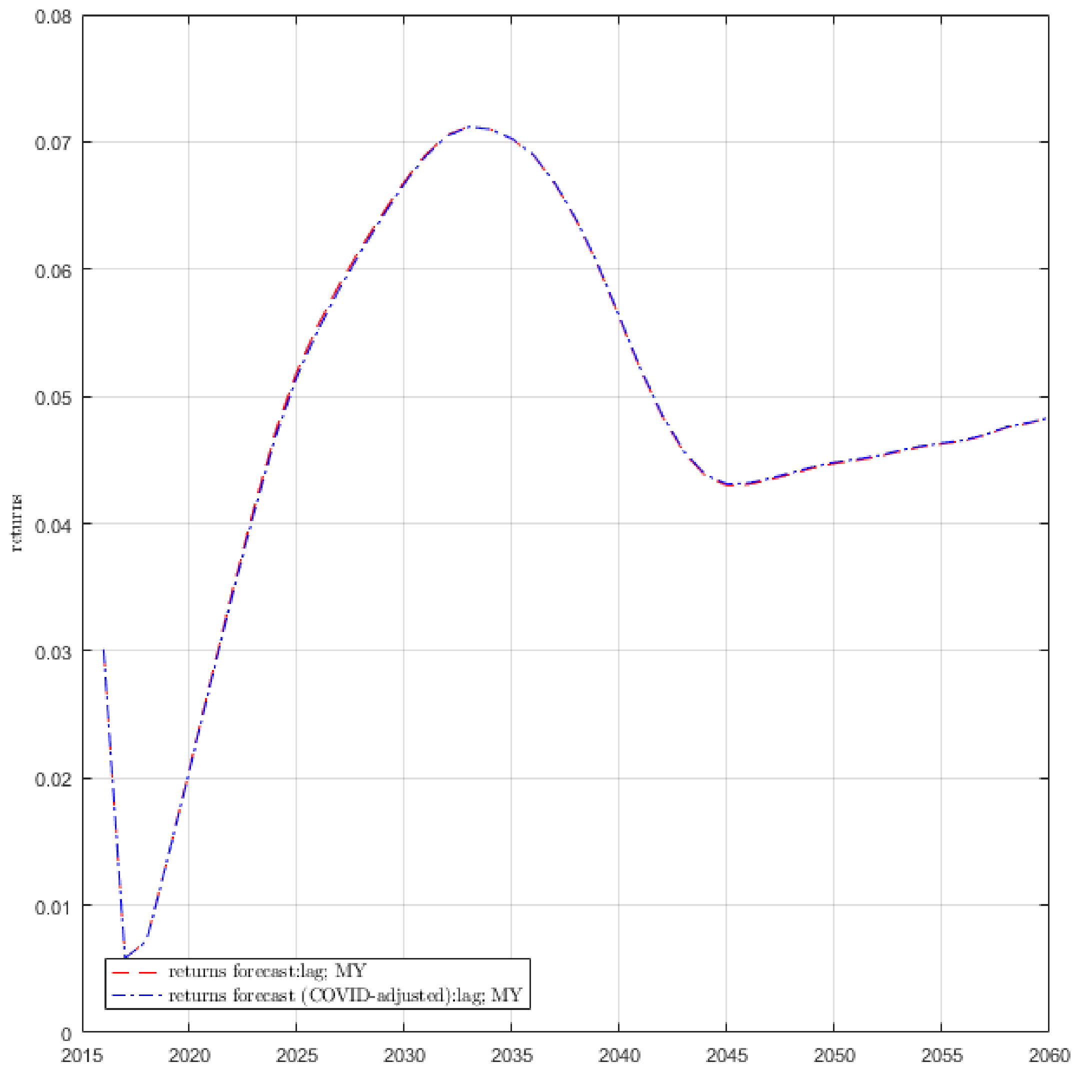

This figure provides MY ratio projections for both unadjusted and Covid-19 adjusted annual projections. The red line presents the MY ratio of the projection from Census Bureau 2017. The blue line shows the MY ratio Covid-19 adjusted projection.

Figure B16: Projections for the dp Ratio Based on Covid19 Adjusted Demographic Ratios, Scenario Three: $\theta = 10\%, t^* = 2030$



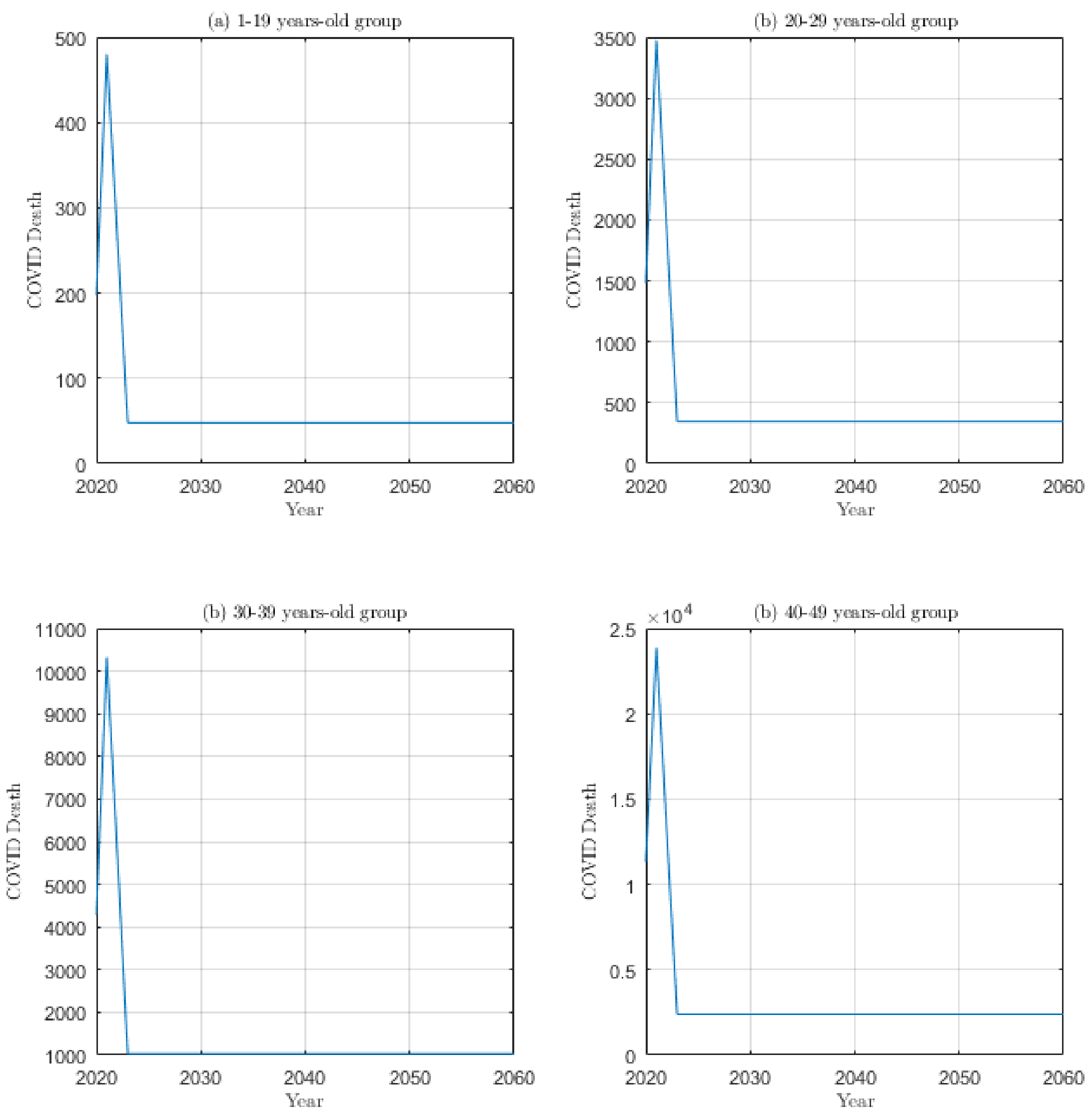

This figure provides forecasts for the dividend-price ratio based on both the Census Bureau projection and the Covid-19 adjusted projection. The red dashed line shows the projection based on the Census Bureau projection. The blue dashed line gives the dp ratio projection from the Covid-19 adjusted demographic projection.

Figure B17: Covid19 Adjusted Projected Five-Year Rolling Average Stock Returns, Scenario Three: $\theta = 10\%, t^* = 2030$



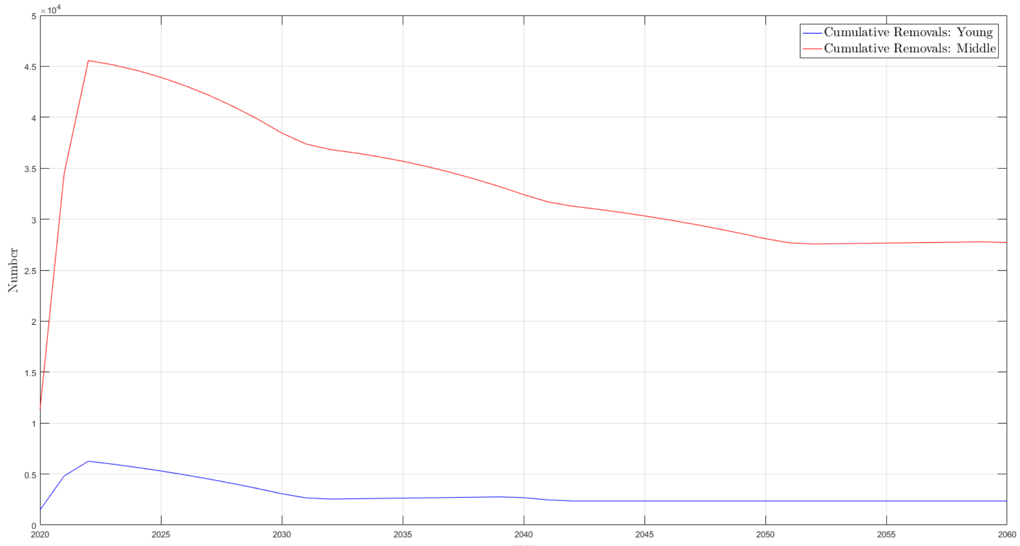

This figure provides forecasts for the five year rolling average return (including dividends) on the S&P 500 index using both the Census Bureau demographic projections and the Covid-19 adjusted demographic projection. The red dashed line shows the projection based on the Census Bureau projection. The blue dashed line gives the projection based on the Covid-19 adjusted demographic projection.

Figure B18: COVID-19 Annual Death: Actual and Projection for Each Age-Group, Scenario One: $\theta = 10\%, t^* = 2023$



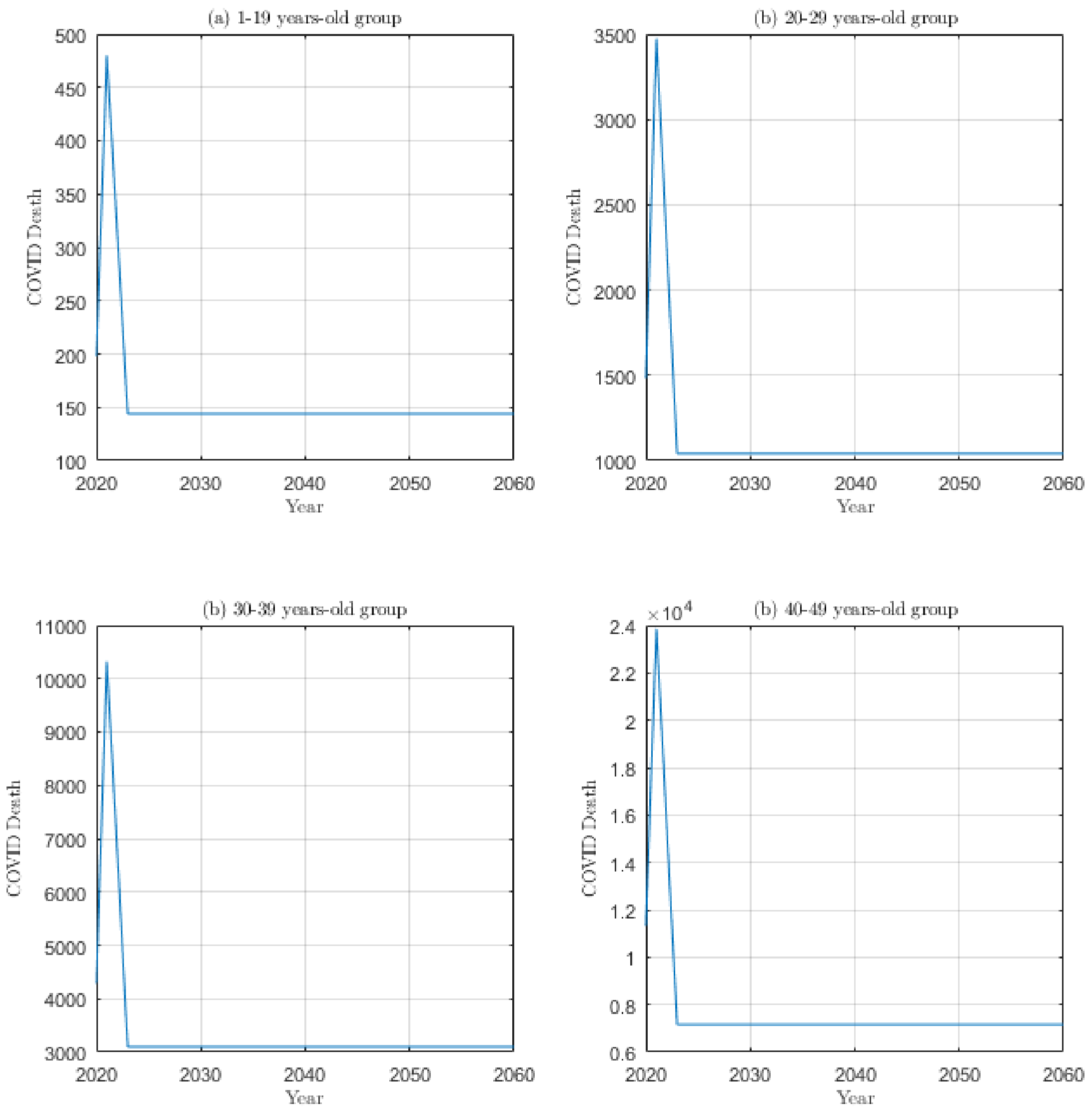

This figure provides actual deaths and annual projections involving COVID-19 for the each age-group. Panel (a), (b), (c), and (d) show the death projection for the children, young, thirty, and middle group, respectively.

Figure B19: Projected Cumulative Removals in Each Year, Scenario One: $\theta = 10\%, t^* = 2023$



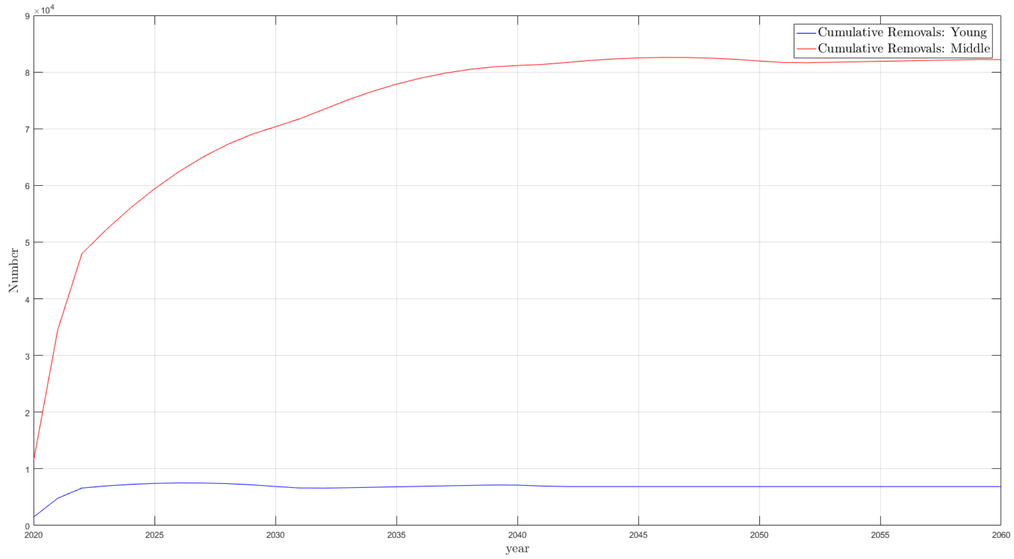

This figure provides the cumulative removals for both young and middle group. The removals are defined in subsection (A3.4). The red line is the projected cumulative removals for the young group in each year. The blue is line shows the removals for the middle group.

Figure B20: COVID-19 Annual Death: Actual and Projection for Each Age-Group, Scenario Two: $\theta = 30\%, t^* = 2023$



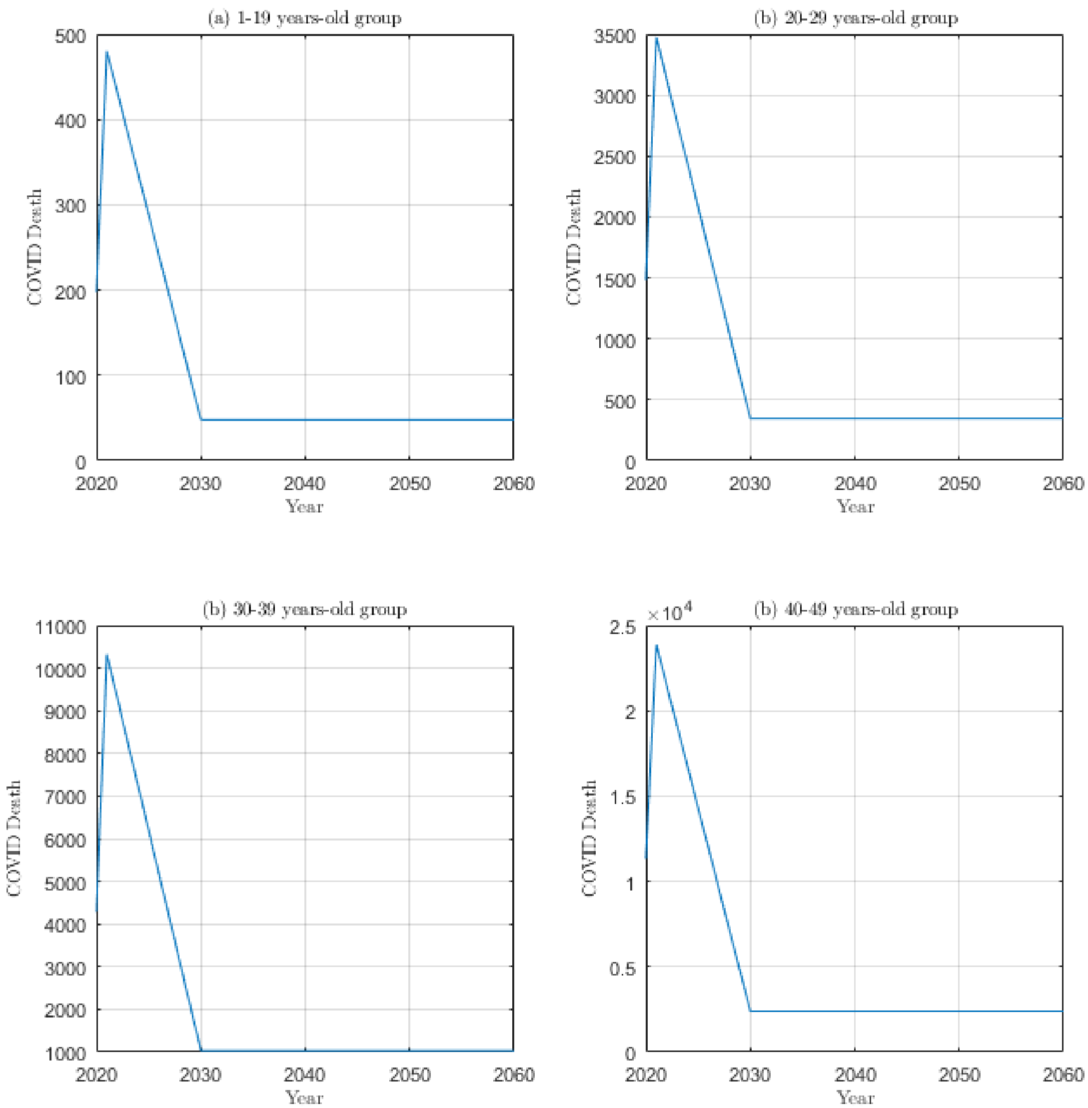

This figure provides actual deaths and annual projections involving COVID-19 for the each age-group. Panel (a), (b), (c), and (d) show the death projection for the children, young, thirty, and middle group, respectively.

Figure B21: Projected Cumulative Removals in Each Year, Scenario Two: $\theta = 30\%, t^* = 2023$



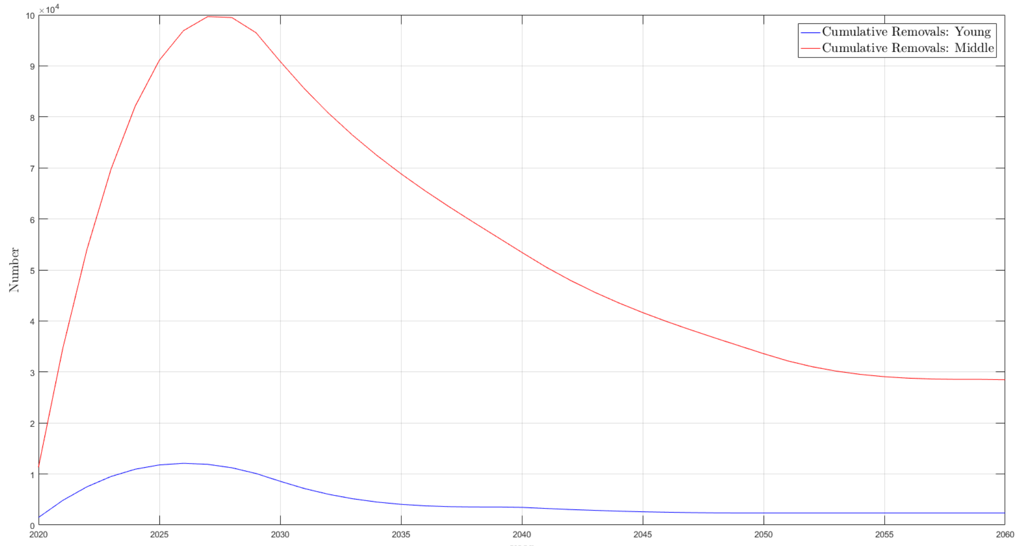

This figure provides the cumulative removals for both young and middle group. The removals are defined in subsection (A3.4). The red line is the projected cumulative removals for the young group in each year. The blue is line shows the removals for the middle group.

Figure B22: COVID-19 Annual Death: Actual and Projection for Each Age-Group, Scenario Three: $\theta = 10\%, t^{*} = 2030$



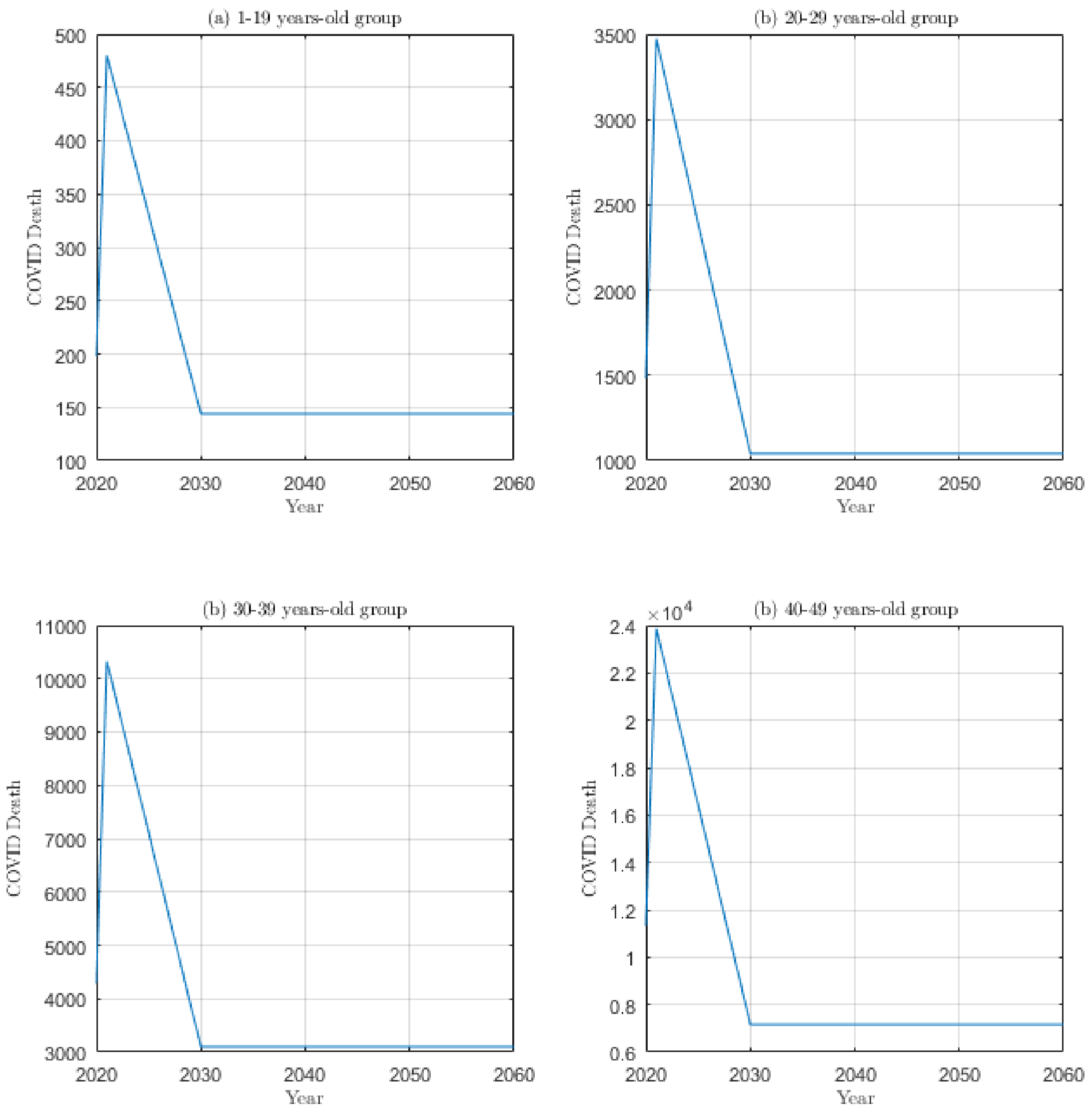

This figure provides actual deaths and annual projections involving COVID-19 for the each age-group. Panel (a), (b), (c), and (d) show the death projection for the children, young, thirty, and middle group, respectively.

Figure B23: Projected Cumulative Removals in Each Year, Scenario Three: $\theta = 10\%, t^* = 2030$



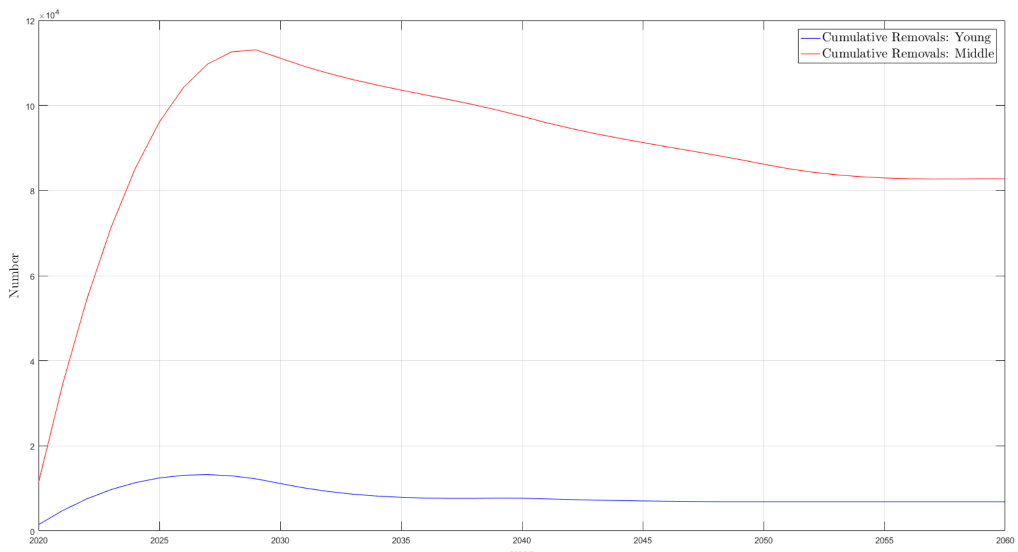

This figure provides the cumulative removals for both young and middle group. The removals are defined in subsection (A3.4). The red line is the projected cumulative removals for the young group in each year. The blue is line shows the removals for the middle group.

Figure B24: COVID-19 Annual Death: Actual and Projection for Each Age-Group, Scenario Four: $\theta = 30\%, t^* = 2030$




This figure provides actual deaths and annual projections involving COVID-19 for the each age-group. Panel (a), (b), (c), and (d) show the death projection for the children, young, thirty, and middle group, respectively.

Figure B25: Projected Cumulative Removals in Each Year, Scenario Four: $\theta = 30\%, t^* = 2030$




This figure provides the cumulative removals for both young and middle group. The removals are defined in subsection (A3.4). The red line is the projected cumulative removals for the young group in each year. The blue is line shows the removals for the middle group.

## References


Campbell, J. Y. and Thompson, S. B. (2007), 'Predicting excess stock returns out of sample: Can anything beat the historical average?', Review of Financial Studies **21**(4), 1509–1531.

Clark, T. E. and McCracken, M. W. (2001), 'Tests of equal forecast accuracy and encompassing for nested models', Journal of Econometrics **105**, 85–110.

Clark, T. E. and McCracken, M. W. (2017), 'Tests of predictive ability for vector autoregressions used for conditional forecasting', Journal of Applied Econometrics **32**(3), 533–553.

Clark, T. E. and West, K. D. (2007), 'Approximately normal tests for equal predictive accuracy in nested models', Journal of Econometrics **138**, 219–311.

Elliott, G., Rothenberg, T. J. and Stock, J. (1996), 'Efficient forecast tests for conditional policy forecasts', Econometrica **64**, 813–836.

Faust, J. and Wright, J. H. (2008), 'Efficient forecast tests for conditional policy forecasts', Journal of Econometrics **146**, 293–303.

Favero, C. A., Gozluklu, A. E. and Tamoni, A. (2011), 'Demographic trends, the dividend-price ratio, and the predictability of long-run stock market returns', Journal of Financial and Quantitative Analysis **46**(05), 1493–1520.

Giacomini, R. and White, H. (2006), 'Tests of conditional predictive ability', Econometrica **74**(6), 1545–1578.

McCracken, M. W. (2007), 'Asymptotics for out of sample tests of granger causality', Journal of Econometrics **140**, 719–752.

Pesavento, E. (2004), 'An analytical evaluation of the power of tests for the absence of cointegration', Journal of Econometrics **122**(2), 375–421.