“Model Complexity, Expectations, and Asset Prices” Online Appendix

POOYA MOLAVI

Northwestern University

ALIREZA TAHBAZ-SALEHI

Northwestern University and CEPR

and

ANDREA VEDOLIN

Boston University, NBER, and CEPR

This online appendix contains four parts. Appendix C presents the proofs and derivations omitted from the main body of the paper, Appendix D contains some additional theoretical results, Appendix E is a data appendix, and Appendix F reproduces the exercises in Subsections 5.2 and 5.3 for each of the currencies in our sample separately.

C. LEMMA FOR THE PROOF OF THEOREM 1

This Appendix contains the statement and proof of a lemma that is used in the proof of Theorem 1.

Lemma C.1. *Let* $\mathbf{M}$ *and* u *be defined as in* (B.11) *and let* $\hat{\sigma}_x^2 = c'\hat{\mathbf{\Sigma}}c$. *Minimizing* (B.13) *with respect to* $(\mathbf{M}, u, \hat{\sigma}_x^2)$ *and subject to the constraint that* $\rho(\mathbf{M}(\mathbf{I}-uu')\mathbf{M}')<1$ *is equivalent to minimizing* (B.10) *with respect to* $(\mathbf{A}, \mathbf{B}, c)$.

Proof. In the proof of Theorem 1, we showed that for any $(\mathbf{A}, \mathbf{B}, c)$, there exists $(\mathbf{M}, u, \hat{\sigma}_x^2)$ such that the values of (B.10) and (B.13) coincide. Furthermore, (B.11) implies that $\mathbf{M}$ and u satisfy $\rho(\mathbf{M}(\mathbf{I}-uu')\mathbf{M}')<1$. To see this, observe that plugging in for $\mathbf{A}$ and c in (B.7) in terms of $\mathbf{M}$ and u in (B.11) implies that $\hat{\mathbf{\Sigma}} = \hat{\mathbf{\Sigma}}^{1/2}\mathbf{M}(\mathbf{I}-uu')\mathbf{M}'\hat{\mathbf{\Sigma}}^{1/2} + \mathbf{B}\mathbf{B}'$. As a result,

$$\mathbf{I} - \mathbf{M}(\mathbf{I}-uu')\mathbf{M}' = \hat{\mathbf{\Sigma}}^{-1/2}\mathbf{B}\mathbf{B}'\hat{\mathbf{\Sigma}}^{-1/2}.$$

Since the right-hand side of the above equation is positive-definite, it follows that $\rho(\mathbf{M}(\mathbf{I}-uu')\mathbf{M}')<1$.

Next, we show that for any $(\mathbf{M}, u, \hat{\sigma}_x^2)$ such that $\rho(\mathbf{M}(\mathbf{I}-uu')\mathbf{M}')<1$, there always exists $(\mathbf{A}, \mathbf{B}, c)$ such that the values of (B.10) and (B.13) coincide. To this end, let $(\mathbf{M}, u, \hat{\sigma}_x^2)$ be an arbitrary tuple such that $\rho(\mathbf{M}(\mathbf{I}-uu')\mathbf{M}')<1$ and define $(\mathbf{A}, \mathbf{B}, c)$ as follows:

$$\mathbf{A}=\mathbf{M}, \qquad \mathbf{B}=\hat{\sigma}_x(\mathbf{I}-\mathbf{M}(\mathbf{I}-uu')\mathbf{M}')^{1/2}, \qquad c=u.$$

Given these definitions, it is immediate to verify that the solution to the algebraic Riccati equation (B.7) and the corresponding Kalman gain in (B.8) are given by $\hat{\mathbf{\Sigma}}=\hat{\sigma}_x^2\mathbf{I}$ and $g=\mathbf{M}u$, respectively. This also implies that $\mathbf{M}(\mathbf{I}-uu')=\mathbf{A}-gc'$. Replacing for $\mathbf{M}(\mathbf{I}-uu')$, $\mathbf{M}u$, and u' on the right-hand side of (B.13) by, respectively, $\mathbf{A}-gc'$, g, and c' leads to the right-hand side of (B.10). ‖

D. ADDITIONAL THEORETICAL RESULTS

D.1. *Over- and Under-Reaction to Information*

Recall from Proposition 3 that when the set of models entertained by agents is not rich enough to contain the true data-generating process (i.e., when $k<n$), they end up with subjective expectations that generate return predictability. This departure from rational expectations is a consequence of the fact that, due to their misspecified model of the data-generating process, agents do not incorporate new pieces of information into their forecasts as they would have had they known the true process.

We show that whether agents over- or under-react to new information cannot be decoupled from the environment they live in. In particular, we argue that agents' forecasts may exhibit systematic over- or under-reaction to news depending on (i) the statistical properties of the underlying data-generating process and (ii) the horizon of interest. This means that, in our framework, over- and under-reaction of expectations are endogenous and are not baked into agents' expectations formation process.

To measure the extent of over- and under-reaction to new information at different horizons, we follow Coibion and Gorodnichenko (2015) and consider the family of regressions

$$x_{t+h}-\mathbb{E}_t[x_{t+h}]=\alpha_h^{\mathrm{CG}}+\beta_h^{\mathrm{CG}}(\mathbb{E}_t[x_{t+h}]-\mathbb{E}_{t-1}[x_{t+h}]), \tag{D.1}$$

where, as before, x_t is the realization of the fundamental at time t and $\mathbb{E}[\cdot]$ denotes agents' subjective expectations. The regressand on left-hand side of (D.1) is agents' forecast error for realization of the fundamental h periods in the future, while the regressor on the right-hand side is their latest forecast revision. Thus, $\beta_h^{\mathrm{CG}}>0$ means that when agents revise their h-step-ahead forecast of the fundamental upward at time t, they tend to undershoot its eventual realization at time $t+h$. In other words, agents systematically under-react to new information at time t. By a similar logic, $\beta_h^{\mathrm{CG}}<0$ means that agents tend to systematically over-react to new information when forming expectations about the realization of the fundamental h periods in the future. As is well known, under rational expectations (i.e., when $\mathbb{E}[\cdot]=\mathbb{E}^*[\cdot]$), forecast revisions are orthogonal to future forecast errors, which means that $\beta_h^{\mathrm{CG}}=0$ for all horizons $h\geq 1$.

Turning to our framework, suppose that agents are constrained to using single-factor models (i.e., $k=1$) to make forecasts about the future realization of the fundamental, even though the true data-generating process may be driven by $n\neq k$ factors. Under such a specification, the slope coefficient of regression (D.1) at horizon h is given by

$$\beta_h^{\mathrm{CG}}=\frac{\xi_h^*-\xi_1^*\xi_{h+1}^*}{\xi_1^{*h}(1-\xi_1^{*2})}-1, \tag{D.2}$$

where ξ_h^* denotes the autocorrelation of the fundamental at lag h.

Two observations are immediate. First, whether agents over- or under-react to new information depends on the statistical properties of the underlying data-generating process. For example, it is straightforward to verify that $\beta_1^{\text{CG}} > 0$ if $\xi_1^{*2} > \xi_2^*$, whereas $\beta_1^{\text{CG}} < 0$ if $\xi_1^{*2} < \xi_2^*$. Thus, the same agent with the same level of sophistication may end up with forecasts that under- or overshoot the eventual realization of the fundamental depending on the environment. Second, holding the underlying data-generating process fixed, the sign and the magnitude of the right-hand side of (D.2) varies with h. This means that agents' expectations may over-react to new information at some horizon, while simultaneously under-reacting at others. As already discussed, the exact pattern of over- and under-reaction depends on the process that governs the fundamental.

Derivation of Equation (D.2).. Since $k=1$, agents' model can be represented by the tuple of scalars $\theta=(a,b,c)$, where $a\in(-1,1)$ denotes the persistence parameter of the underlying factor. This means that $\mathbb{E}_t[x_{t+h}]=a^h x_t$ for all $h\geq 0$. As a result, the slope coefficient of the regression in (D.1) is given by

$$\beta_h^{\text{CG}} = \frac{\mathbb{E}^*[(x_{t+h}-a^h x_t)(x_t - a x_{t-1})]}{a^h \mathbb{E}^*[(x_t - a x_{t-1})^2]} = \frac{\xi_h^* - a\xi_{h+1}^* - a^h(1-a\xi_1^*)}{a^h(1+a^2-2a\xi_1^*)}, \tag{D.3}$$

where ξ_h^* denotes the autocorrelation of the fundamental at lag h. Next, recall from the proof of Proposition 6 that when agents are restricted to single-factor models, the objective function in (12) reduces to (B.23), in which $|m|<1$ is a scalar. Optimizing (B.23) over m implies that $m=\xi_1^*$. Therefore, by equation (B.11), the persistence parameter is given by $a=\xi_1^*$. Plugging this expression into (D.3) establishes (D.2).

D.2. *Micro-Founded Model for Heterogenous-Agent Economy*

As discussed in Appendix A.3, Proposition A.3 applies to any asset pricing framework that has a reduced-form representation in the form of equation (A.3). In this appendix, we present a stylized micro-founded model with such a property. The model, which is a variant of the heterogenous-agent economy of Allen, Morris, and Shin (2006), has a reduced-form representation identical to equation (A.3) and reduces to equation (2) when all agents are identical.[1]

Consider a discrete-time economy with a single consumption good and a single risky asset. The risky asset is in zero net supply and delivers a dividend stream $\{x_t\}_{t=-\infty}^{\infty}$ of the consumption good. The economy is populated by overlapping generations of agents, who live for 2 periods. A new generation of agents of unit mass is born at each date t. Agents in the same generation have identical preferences and can observe the entire past realizations of x_t but may have heterogenous subjective expectations about the process that drives the asset's dividend stream (and hence, have different expectations about future dividends). Specifically, we assume that a fraction $1-\gamma$ of agents are restricted to using models consisting of at most k factors, while the remaining γ fraction are unconstrained and can entertain models with any number of factors.

All agents only consume when they are old. When young, agents build up a position in the asset, which they then unwind in the next period—when they are old—to acquire the

1. Also see Bacchetta and Van Wincoop (2006) for a similar model.

consumption good. The utility of an agent $i \in [0,1]$ who is born in period t and acquires q_{it} units of the asset is given by

$$u_i(q_{it}) = q_{it}(x_t - y_t + y_{t+1}) - \frac{1}{2}\gamma q_{it}^2,$$

where y_t denotes the price of the asset at time t in units of the consumption good and $\gamma > 0$ parameterizes a quadratic cost of trading the asset.[2]

Given the above, it is immediate that the demand for the asset by an agent i born at time t is given by $q_{it} = \frac{1}{\gamma}(x_t - y_t + \mathbb{E}_{it}[y_{t+1}])$, where $\mathbb{E}_{it}[\cdot]$ denotes the agent's subjective expectations. Consequently, market clearing implies that the price of the asset at time t satisfies

$$y_t = x_t + \int_0^1 \mathbb{E}_{it}[y_{t+1}]\mathrm{d}i.$$

In other words, the price of the asset at time t is equal to its dividend plus the cross-sectional average of agents' subjective expectations of the price the next period. Given our assumption that only a fraction $1-\gamma$ of the agents are subject to our behavioral constraint, this implies that $y_t = x_t + \bar{\mathbb{E}}_t[y_{t+1}]$, where $\bar{\mathbb{E}}[\cdot] = \gamma \mathbb{E}^*[\cdot] + (1-\gamma)\mathbb{E}[\cdot]$.

E. DATA APPENDIX

In this appendix, we provide a brief description of how we construct the realized forecast errors of interest rate differentials used in the analysis in Subsection 5.3.

We obtain interest rate forecasts from *Consensus Economics* between June 1998 and December 2021 for the following set of countries: Canada, the Eurozone, Japan, Switzerland, the United Kingdom, and the United States. The surveys query respondents every month about a number of country-specific macroeconomic and financial variables, including 3-month-ahead forecasts of 3-month interest rates. Using these surveys, we construct a "consensus forecast" by calculating the median forecast for each country. This results in a monthly sample for 3-month-ahead consensus forecasts of 3-month interest rates in each of the six countries. We restrict our sample to the June 1998–December 2021 period in order to have a non-trivial cross-section of different forecasts in each country.

To measure the realized counterparts to the survey forecasts, we obtain the following 3-month interest rates:

- Canada: 3-Month Treasury Bill rate available from Bank of Canada webpage;
- Eurozone: 3-Month or 90-day Rates and Yields: Interbank Rates available from FRED;
- Japan: 3-Month Yen Certificates of Deposit until June 2010; 3-month Yen TIBOR Rate from July 2010, available from the Bank of Japan webpage;
- Switzerland: 3-Month or 90-day Rates and Yields: Eurodollar Deposits available from FRED;

2. As in Allen, Morris, and Shin (2006), the assumption of overlapping generations of agents is intended to abstract from dynamic incentives. The quadratic trading costs are imposed to ensure that disagreement among agents does not lead to unbounded trade in equilibrium.

- United Kingdom: 3-Month or 90-day Rates and Yields: Interbank Rates available from FRED;
- United States: 3-Month Treasury Bill Secondary Market Rate available from FRED.

We then construct the forecast errors for each interest rate by subtracting the survey forecasts from their realized counterparts. We follow a similar procedure to construct forecast errors for interest rate differentials.

F. CROSS-SECTIONAL EVIDENCE

In the application on UIP violations in Section 5, we followed Engel (2016) and focused on trade-weighted average exchange rates and interest rate differentials of various currencies vis-à-vis the United States. In this appendix, we redo the exercises in Subsections 5.2 and 5.3 for each of the currencies in our sample separately.

F.1. *Return Predictability*

We calculate the model-implied slope coefficients of the return predictability regression (6) for each currency separately (all against the U.S. dollar) and plot them against the corresponding coefficients estimated from the data. Figure F.1 depicts the results for $k=1$ at different horizons. Recall from Theorem 1 and Proposition 6 that model-implied slope coefficients depend on the shape of the autocorrelation function of the fundamental (in this case, the interest rate differential between the corresponding country and the U.S.). Therefore, as the autocorrelations differ in the cross-section of countries, so should the model-implied slope coefficients, β_h^{rx}. Nonetheless, as Figure F.1 illustrates, the patterns of model-implied slope coefficients closely track those of the coefficients estimated from return data for each currency. This relationship holds both at short horizons when most coefficients are positive, as well as for the longer horizons. Once again, we emphasize that we do not use exchange rate and return data for generating any of the model-implied slope coefficients. Finally, note that, as in Figure 2 for the trade-weighted basket of currencies, the model-implied coefficients in Figure F.1 underestimate the magnitudes of their empirical counterparts.

F.2. *Forecast-Error Predictability*

We also reproduce the results on forecast-error predictability in Subsection 5.3 separately for each of the countries in our sample. To estimate the coefficients of forecast-error predictability regressions, we regress the forecast error of each interest rate differential relative to the U.S. (constructed from *Consensus Economics* data) on the corresponding realized interest rate differential. We also obtain the model-implied slope coefficients of the forecast-error predictability regression for each country from the corresponding autocorrelation of the interest rate differential process under the assumption that $k=1$ using equation (16). Once again, note that, as the autocorrelations differ in the cross-section of countries, so should the model-implied slope coefficients, β_h^{fe}. Figure F.2 depicts the results.

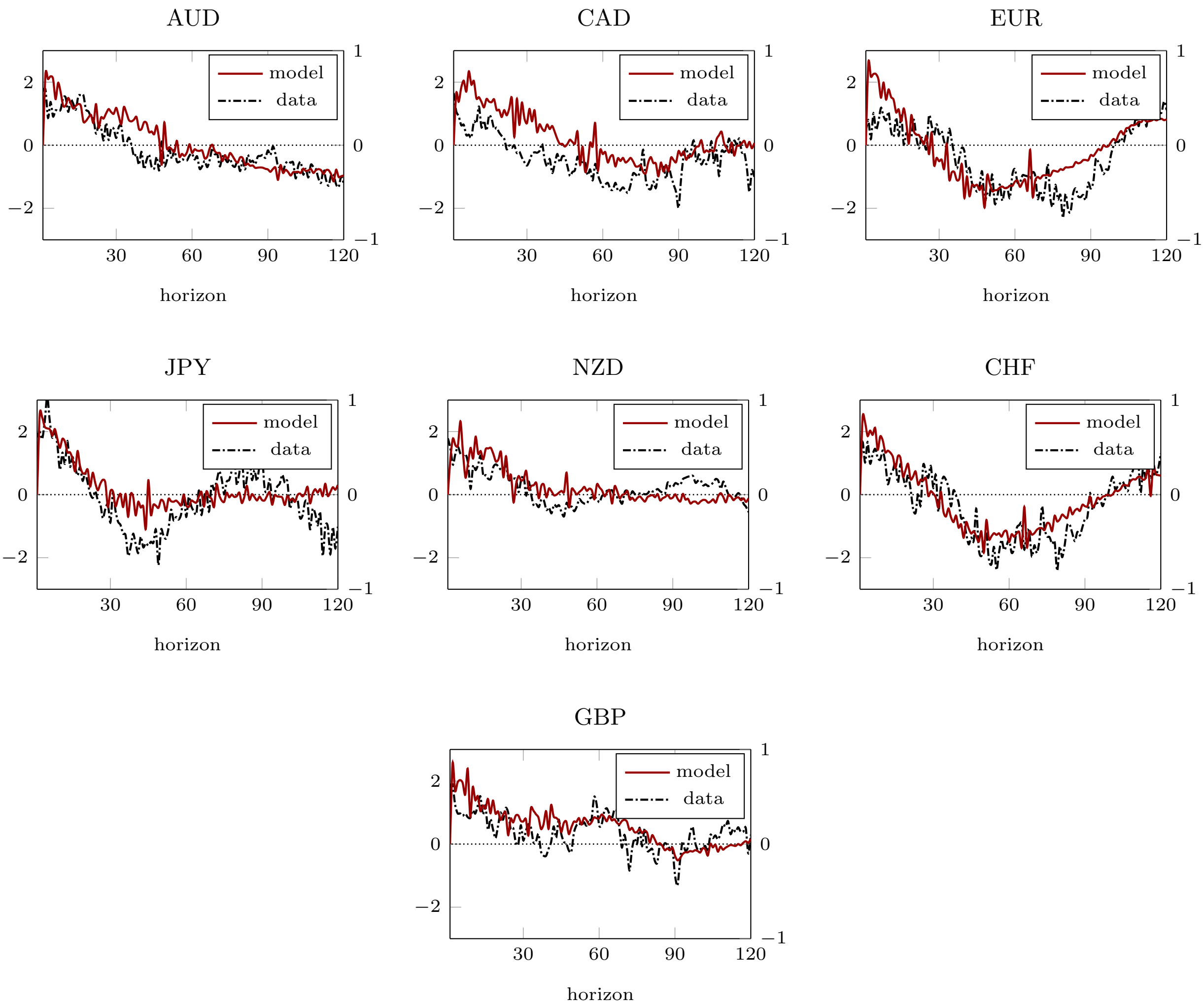

FIGURE F.1
Return Predictability in the Cross-Section

Notes: This figure plots estimated slope coefficients of return predictability regression (left axis) together with model-implied coefficients from a one-factor model given in Proposition 6 (right axis) for each currency separately. Monthly data from January 1985 to December 2021.

F.3. *Testing for k*

Finally, we use our result in Proposition 8 to directly test for the value of k for each of the countries in our sample separately. We follow the same 3-step procedure laid out in Subsection 5.3. For each country, we (i) regress the consensus (3-month-ahead) forecasts of interest rate differential vis-à-vis the United States obtained from survey data on the past realizations of the (one-month CIP-implied) interest rate differential to estimate coefficients $(\phi_1^{(3)}, \phi_2^{(3)}, \ldots)$ in (17); (ii) construct the Hankel matrix $\mathbf{\Phi}$ with typical element $\mathbf{\Phi}_{ij} = \phi_{i+j-1}^{(3)}$; and (iii) test for the rank of $\mathbf{\Phi}$. Note that, technically speaking,

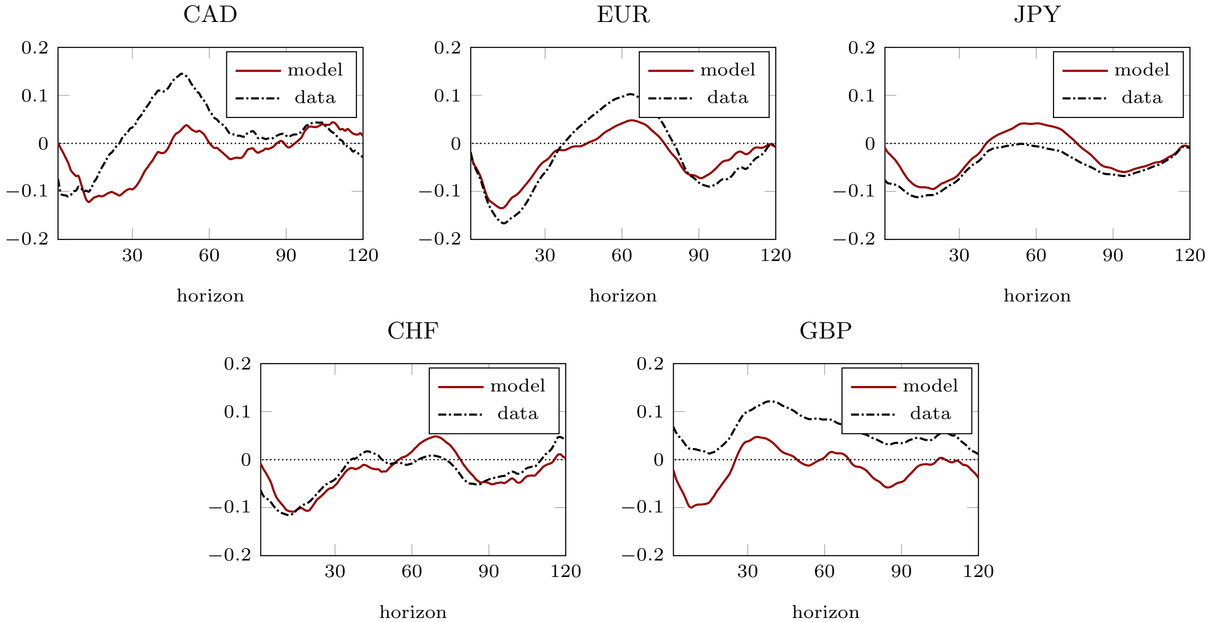

FIGURE F.2

Forecast-Error Predictability in the Cross-Section

Notes: This figure plots estimated slope coefficients of forecast-error predictability regression of the interest rate differential process (vis-à-vis the U.S.) for each country together with the corresponding model-implied coefficients from a one-factor model. Monthly data from June 1998 to December 2021.

implementing step (i) requires regressing the forecasts of interest rate differentials on all past realizations of interest rate differentials. To implement this step empirically, we perform the above procedure while varying the number of lags used for estimating (17) from 10 to 20. Additionally, note that step (iii) of the estimation procedure requires testing for the rank of an infinite-dimensional matrix. This, however, is not an issue for our empirical implementation, as by Kronecker's Theorem (Al'pin, 2017, Theorem 2), rank($\mathbf{\Phi}$) is equal to the rank of the finite-dimensional Hankel matrix constructed from coefficients $(\phi_1^{(3)},\ldots,\phi_s^{(3)})$ for any odd integer s such that $s \geq 2k-1$. In what follows, we set $s=7$, but find the same result for other values of s.

Table F.1 reports the p-values corresponding to the rank test of Donald, Fortuna, and Pipiras (2007) for the null hypothesis that rank($\mathbf{\Phi}$)$=1$ for different countries (including the trade-weighted average studied in Subsection 5.3) and different lags in equation (17).

REFERENCES

Allen, Franklin, Stephen Morris, and Hyun Song Shin (2006), "Beauty contests and iterated expectations in asset markets." *Review of Financial Studies*, 19(3), 719–752.

Al'pin, Yu. A. (2017), "The Hankel matrix rank theorem revisited." *Linear Algebra and its Applications*, 534, 97–101.

Bacchetta, Philippe and Eric Van Wincoop (2006), "Can information heterogeneity explain the exchange rate determination puzzle?" *American Economic Review*, 96(3), 552–576.

TABLE F.1
Testing for k

	Country					
No. of lags	CAD	EUR	JPY	CHF	GBP	Average
10	0.60	0.94	0.88	0.92	0.18	0.90
11	0.59	0.95	0.18	0.94	0.18	0.95
12	0.55	0.97	0.33	0.69	0.16	0.92
13	0.42	0.97	0.62	0.00	0.18	0.78
14	0.29	0.98	0.61	0.94	0.10	0.95
15	0.27	0.96	0.58	0.91	0.07	0.95
16	0.10	0.98	0.65	0.97	0.07	0.94
17	0.08	0.98	0.73	0.98	0.07	0.96
18	0.08	0.97	0.71	0.88	0.08	0.95
19	0.09	0.97	0.66	0.84	0.10	0.95
20	0.14	0.98	0.65	0.79	0.09	0.95

Notes: This table reports p-values of the rank test of Donald, Fortuna, and Pipiras (2007) for the null hypothesis that $\text{rank}(\boldsymbol{\Phi}) = 1$ for different countries. Matrix $\boldsymbol{\Phi}$ is a Hankel matrix constructed from coefficients in (17). The first column contains the number of lags on the right-hand side of (17).

Coibion, Olivier and Yuriy Gorodnichenko (2015), "Information rigidity and the expectations formation process: A simple framework and new facts." *American Economic Review*, 105(8), 2644–2678.

Donald, Stephen G., Natércia Fortuna, and Vladas Pipiras (2007), "On rank estimation in symmetric matrices: The case of indefinite matrix estimators." *Econometric Theory*, 23(6), 1217–1232.

Engel, Charles (2016), "Exchange rates, interest rates, and the risk premium." *American Economic Review*, 106(2), 436–474.