Supplementary Material for

Multilateral Political Effects on Outbound Tourism

1 Derivation of the Empirical Model from the Theoretical Foundation

We explicitly incorporate the three mechanisms of political effects (country image, nationalism, and government intervention) into a gravity-like tourism demand model. The gravity model was initially transplanted into international trade literature (McCallum, 1995; Anderson & Van Wincoop, 2003) and later transplanted to international tourism literature (Kimura & Lee, 2006; Keum, 2010). In the traditional bilateral gravity model, tourist flow (F_{od}) from origin country o to destination country d is determined by the "masses" (M) of the two countries (usually GDP per capita) and the bilateral distance (D_{od}), multiplied by a constant G. A log-linear transformation is usually used in empirical implementation:

$$F_{od} = G \times \frac{M_o^{\beta_o} M_d^{\beta_d}}{D_{od}^{\beta_{od}}} \rightarrow \ln F_{od} = intercept + \boldsymbol{\beta}'[\ln M_o, \ln M_d, \ln D_{od}]. \quad (1)$$

There are essentially two types of factors on the right-hand side of equation (1): terms related to only one country (M_o, M_d) and terms related to country pair (D_{od}). Morley et al. (2014) structurally derive this gravity equation from microeconomic behavior (known as "micro-foundation revolution" in macroeconomic literature). We extend their bilateral and static framework to allow for both *multilateral interdependence* and *intertemporal dynamics*. The former is derived theoretically, and the latter is formulated empirically.

1.1 The Theoretical Model with Multilateral Interdependence

An individual (indexed by i) from an origin country maximizes his/her utility $U(\cdot)$ by choosing the numbers of visit, $f_d(i)$, to a set of destination countries $d \in [1, N]$ under an optimally allocated tourism budget, which is a power function of his/her total income $m(i)$—in other words, we assume a constant income elasticity (μ) of demand for tourism.

$$\max_{\{f_d(i)\}_{d=1}^N} U(f_1(i), \dots, f_d(i), \dots, f_N(i)|\boldsymbol{\alpha}), \text{ subject to: } \sum_{d=1}^N c_d f_d(i) = m(i)^\mu. \quad (2)$$

In the objective function, $\boldsymbol{\alpha}$ is a vector of parameters describing the preferences $U(\cdot)$ over N tourism destinations. In the budget constraint, c_d denotes the cost of traveling to destination d, which is affected by exchange rate, oil price, national holidays and particularly **political relation**. This setup includes the bilateral model as a special case where the decisions on visiting any two destinations are totally independent. In the bilateral model, the objective function $U(\cdot)$ only includes the visit to one specific destination $f_d(i)$, while visits to the other destinations are absorbed into the "consumption of other goods" (Morley et al., 2014).

If interior solution exists, the optimality conditions of the maximization problem (2) lead to a standard micro-economic principle—the marginal rate of substitution (MRS) between any two destinations ($d = J, K$) must equal the cost ratio of the two:

$$MRS_{JK} \equiv \frac{\frac{\partial U}{\partial f_J(i)}}{\frac{\partial U}{\partial f_K(i)}} = \frac{c_J}{c_K}. \quad (3)$$

Without losing generality, consider a minimalist multilateral model with three countries (a trilateral model) to illustrate the interdependence feature and to fit the empirical context of this paper. Assume the origin country (China) has two alternative destination countries J (Japan) and K (Korea). To arrive at an analytical solution, we specify a Cobb-Douglas utility for the objective function in (2), but the conclusions and implications are generalizable to any constant-elasticity-of-substitution preferences class. The substitutability between the two destinations is shaped by **nationalism** of the tourists—a stronger nationalist sentiment means a greater sensitivity switching between alternative destinations. The parameters α_J and α_K are the *relative* attractiveness or utility weights of Japan and Korea, reflecting the **country image** of the two countries.

$$U(\cdot) \equiv f_J(i)^{\alpha_J} f_K(i)^{\alpha_K}, \text{ where } \alpha_J + \alpha_K = 1. \quad (4)$$

Intuitively, the *relative* attractiveness (α_d, where $d = J, K$) is defined by *absolute* attractiveness (A_d), which describes the degree of attraction of country $d = J, K$ as tourism destination to the origin country. By definition, the relationship between absolute and relative attractiveness is:

$$\alpha_d \equiv \frac{A_d}{A_J + A_K}, \text{ where } A_d > 0. \quad (5)$$

Furthermore, absolute attractiveness A_d (defined to be positive to ensure $0 < \alpha_d < 1$) depends on the natural, cultural, institutional, economic and political characteristics. Given that the paces of change in natural, cultural and institutional characteristics are much slower than that of monthly tourist flows (the dependent variable), we can treat those factors as constant and include them in the intercept. To explain the *monthly* variations, more relevant factors are the economic and political covariates which share a similar frequency of change, such as GDP per capita ($GDPP_d$), events/disasters in the destination (E_d, D_d) and political relation (PR_d) in light of the literature on **country image** and **nationalism**. Following the convention in the empirical literature, a multiplicative power-exponential function $A(\cdot)$ is used to parameterize A_d, so that the coefficients can be interpreted as elasticities (for continuous variables) or quasi-elasticities (for discrete and dummy variables):

$$A_d = A(GDPP_d, E_d, D_d, PR_d), \text{ where } \frac{\partial A_d}{\partial PR_d} > 0. \quad (6)$$

In addition to the effect on travel preferences, political relations also affect travel costs as discussed in the literature on **government intervention**. Therefore, apart from the traditional factors that affect the travel cost ($\mathbf{z}$: exchange rate, oil price and holiday), we also allow political relation to have an impact on c_d in a multiplicative power-exponential function $c(\cdot)$:

$$c_d = c(\mathbf{z}_d, PR_d), \text{ where } \frac{\partial c_d}{\partial PR_d} < 0. \tag{7}$$

Substitute specifications (4)-(7) into (3) and combine with the budget constraint in (2), resulting in the optimal $f_d(i)^*$ for individual i and the aggregate flows F_d:

$$f_d(i)^* = \frac{A_d}{A_J + A_K} \frac{m(i)^\mu}{c_d} = \frac{A(GDPP_d, \ldots, PR_d)}{A(GDPP_J, \ldots, PR_J) + A(GDPP_K, \ldots, PR_K)} \frac{m(i)^\mu}{c(\mathbf{z}_d, PR_d)}. \tag{8}$$

$$F_d = \sum_i f_d(i)^* = \frac{A_d}{A_J + A_K} \frac{\sum_i m(i)^\mu}{c_d} = \frac{A(GDPP_d, \ldots, PR_d)}{A(GDPP_J, \ldots, PR_J) + A(GDPP_K, \ldots, PR_K)} \frac{\sum_i m(i)^\mu}{c(\mathbf{z}_d, PR_d)}. \tag{9}$$

We can immediately spot the interdependence feature from equation (9), as factors related to Japan and Korea enter the solutions for both flows. In other words, what happens in Korea influences China-Japan tourism and vice versa. Let us focus on the effect of China-Japan political relations to derive the properties of this multilateral model, since the conclusions for China-Korea political relations are congruent.

The partial derivative of equation (9) in the instance of Japan analytically demonstrates the **bilateral effect** of China-Japan political relation on China-Japan flow (F_J). It is worth noting that, in our model, there are two reinforcing paths via which political relation can affect the tourist flow: one is by the subjectively perceived attractiveness of the destination (the first term), and the other is by the objectively incurred costs of travel (the second term). As shown in equation (10), both paths are positive.

$$\frac{\partial F_J}{\partial PR_J} = \frac{\partial F_J}{\partial A_J} \times \frac{\partial A_J}{\partial PR_J} + \frac{\partial F_J}{\partial c_J} \times \frac{\partial c_J}{\partial PR_J} = \underbrace{\underbrace{\frac{A_K F_J}{A_J(A_J + A_K)}}_{>0} \times \underbrace{\frac{\partial A_J}{\partial PR_J}}_{>0}}_{>0} + \underbrace{\underbrace{\left(-\frac{F_J}{c_J}\right)}_{<0} \times \underbrace{\frac{\partial c_J}{\partial PR_J}}_{<0}}_{>0} > 0 \tag{10}$$

More importantly, we can derive the **multilateral effect** of China-Japan political relation on the China-Korea flow in equation (11). This (negative) interdependence is the novel feature of the multilateral model:

$$\frac{\partial F_K}{\partial PR_J} = \frac{\partial F_K}{\partial A_J} \times \frac{\partial A_J}{\partial PR_J} = \underbrace{\left(-\frac{F_K}{A_J + A_K}\right)}_{<0} \times \underbrace{\frac{\partial A_J}{\partial PR_J}}_{>0} < 0 \tag{11}$$

1.2 The Empirical Model with Intertemporal Dynamics

To get ready for the empirical analysis, we further log-linearize the solution (9) and add an error term $\boldsymbol{\epsilon}_t$ to capture the residuals that cannot be explained by the right-hand-side covariates. These algebraic transformations result in the log-linearized multilateral gravity equation (12). Mathematical details of these steps can be found in the Appendix.

$$\underbrace{\begin{bmatrix}\ln F_{CJ}\\ \ln F_{CK}\end{bmatrix}}_{\equiv\mathbf{y}_t} = \underbrace{\begin{bmatrix}\boldsymbol{\beta}\mathbf{0}'_J & \boldsymbol{\beta}\mathbf{1}'_J & \boldsymbol{\beta}\mathbf{2}'_{CJ}\\ \boldsymbol{\beta}\mathbf{0}'_K & \boldsymbol{\beta}\mathbf{1}'_K & \boldsymbol{\beta}\mathbf{2}'_{CK}\end{bmatrix}}_{\equiv\mathbf{B}} \underbrace{\begin{bmatrix}\mathbf{x0}_t\\ \mathbf{x1}_t\\ \mathbf{x2}_t\end{bmatrix}}_{\equiv\mathbf{x}_t} + \underbrace{\begin{bmatrix}\epsilon_J\\ \epsilon_K\end{bmatrix}}_{\equiv\boldsymbol{\epsilon}_t}, \text{ or just } \mathbf{y}_t = \mathbf{B}\mathbf{x}_t + \boldsymbol{\epsilon}_t. \qquad (12)$$

In equation (12), the deterministic term ($\mathbf{x0}_t$) contains a constant component, a trend component and a seasonality component (monthly dummies). We also include two types of covariates to determine the monthly tourist flows in light of the empirical literature:

(i) the terms related to only one country ($\mathbf{x1}_t$, China, Japan or Korea): GDP per capita, holidays and events/disasters of each country.

(ii) the terms related to country pairs ($\mathbf{x2}_t$, China-Japan or China-Korea): **political relations**, exchange rate and international (Brent) oil price.

Note that (12) includes the traditional bilateral gravity equation (1) as a special case where the coefficients of covariates involving any third country are restricted to zero.

What is still missing from (12) is the distributional properties of the random error $\boldsymbol{\epsilon}_t$, which determines the *dynamic features* of the tourist flow, such as a lag effect (in mean) and conditional heteroscedasticity (in variance). For this purpose, we adopt the MGARCH model (Hoti et al., 2007), which combines the dynamic features of both VAR (Sims, 1980) and GARCH (Engle, 1982). Therefore, two empirical assumptions of the error term $\boldsymbol{\epsilon}_t$ are added to the theoretical model (12):

$$\boldsymbol{\Phi}(L^p)\boldsymbol{\epsilon}_t = \boldsymbol{\Theta}(L^q)\mathbf{u}_t,\ \mathbf{u}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}_t) \qquad (12)\text{a}$$

$$\text{vech}(\boldsymbol{\Sigma}_t) = \mathbf{s}(\text{vech}(\boldsymbol{\Sigma}_{t-1}), \dots; \text{vech}(\mathbf{u}_{t-1}\mathbf{u}'_{t-1}), \dots) \qquad (12)\text{b}$$

In (12)a, $\boldsymbol{\Phi}(L^p)$ and $\boldsymbol{\Theta}(L^q)$ are polynomials of lag operator L of order p and q, while $\mathbf{u}_t$ is the independently and identically distributed random processes disturbing the system. This assumption equips (12) with multilateral (both within-country and cross-country) lag effects. In (12)b, $\mathbf{s}(\cdot)$ is a multivariate linear function, and $\text{vech}(\cdot)$ function stacks the unique elements that lie on or below the main diagonal in a symmetric matrix into a vector. This assumption harnesses (12) with conditional heteroscedasticity to capture the volatility clustering feature in the monthly tourism data (Santamaria & Filis, 2019).

However, the cost of having both multivariate structure and conditional heteroscedasticity is that the general MGARCH model is too flexible due to excessive free parameters. Therefore, various restrictions on $\mathbf{\Sigma}_t$ have been developed to strike a balance between flexibility and parsimony. We adopt the constant conditional correlation assumption because the pattern of interactions in the China-Japan-Korea geopolitics (and therefore the conditional correlation) is relatively stable in the recent decades.

In summary, compared to the traditional bilateral paradigm, the augmented gravity model developed in this paper has three distinctive features: multilateral interdependence ($\mathbf{x}_t \rightarrow \mathbf{y}_t$), multilateral lag effect ($\mathbf{y}_{t-1} \rightarrow \mathbf{y}_t$), and conditional heteroscedasticity ($\mathbf{\Sigma}_t$).

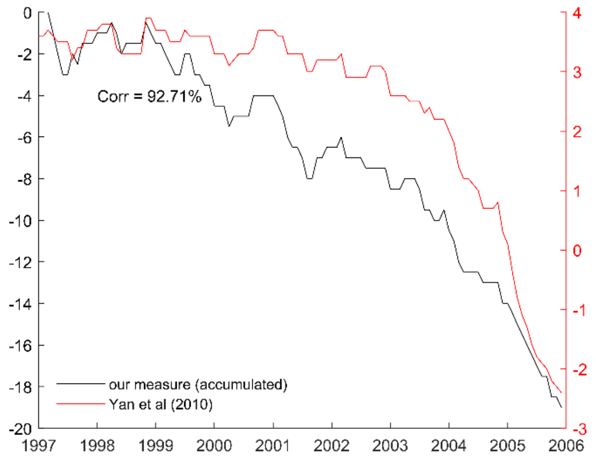
2 Description of the Data

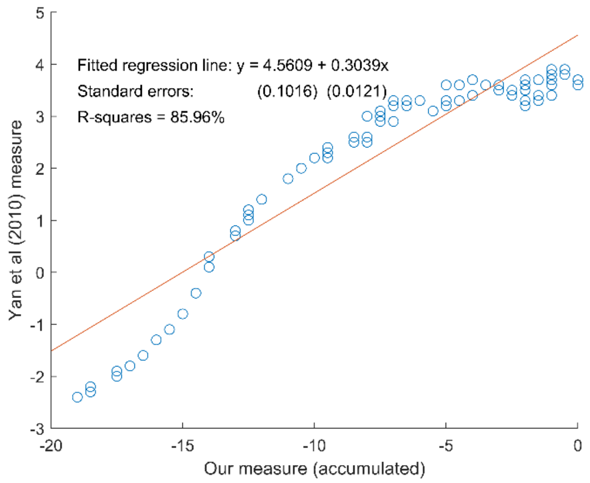
This section describes the data used in the main text in more details, including the comparison with alternative measures of political relations and the diagnostic tests of variables used in the model.

2.1 Alternative Measure of Political Relations

All official statements from the Ministry of Foreign Affairs of the PRC are reported in People's Daily, so we effectively used the same sources of information to quantify political relations as in Yan et al. (2009, 2010) and Du et al. (2017). Moreover, although Yan (2010) and Du et al (2017)'s measurement on political relations may have many merits, we cannot utilize their data for our study because: (i) Their research time period is not up to date—Yan (2010) covers up to 2005, and Du et al (2017) covers up to 2013; (ii) The political relation between China and South Korea, one of our main research objects, is not considered in their research; (iii) Du et al. (2017)' data on the political relations is not available for us.

To show that our measure is consistent with Yan et al (2010), we show the overlapping period of the two measures below (note that Yan et al (2010) is an accumulative measure, so we accumulate our measure to keep comparability). They share the same trend. The correlation coefficient of the two measures are positive and significant (92.71%). We also run a regression between the two measures with a positive and significant coefficient (0.3039) and high R-squares (85.96%). All these evidence shows that our measure is a proper alternative of Yan et al (2010) and their extensions.





2.2 Lists of Events and Disasters

In the regressions, we control for the international events held in the three countries and disasters occurred in the three countries, because they can increase or decrease the travel intentions for tourists. Different events and disasters may have different scales and effects, so the estimated coefficient only describes the average relationship between events/disasters and tourist flows.

Variable		**Year.Month**	**The Lists of Events**
Cdisast	disasters in China	1998.07	1998 China floods
		2003.04	SARS Outbreak
		2008.05	2008 Sichuan earthquake
		2010.04	2010 Yushu earthquake
		2013.03	H7N7 Avian Influenza
		2017.08	2017 Jiuzhaigou earthquake
Jdisast	disasters in Japan	2004.10	2004 Chūetsu earthquake
		2004.12	Typhoon Tokage
		2007.07	2007 Chūetsu offshore earthquake
		2008.06	2008 Iwate earthquake
		2009.08	Tropical Storm Etau
		2011.03	2011 Tōhoku earthquake
		2012.12	2012 Kamaishi earthquake
		2016.04	2016 Kumamoto earthquakes
		2017.07	Tropical Storm Nanmadol
		2018.06-07	2018 Japan floods
Kdisast	disasters in Korea	2002.08	Typhoon Rusa
		2003.09	Typhoon Maemi
		2006.10	North Korea's 1st Nuclear Test
		2009.04	North Korea's 2nd Nuclear Test
		2010.11	The Bombardment of Yeonpyeong
		2011.07	2011 Seoul floods
		2012.12	North Korea's 3rd Nuclear Test
		2016.01	North Korea's 4th Nuclear Test
		2016.09	North Korea's 5th Nuclear Test
		2017.09	North Korea's 6th Nuclear Test
Cevent	events in China	1999.04-10	Expo 1999 Kunming
		2001.08-09	2001 Summer Universiade
		2008.08	Beijing 2008 Summer Olympics
		2009.02	2009 Winter Universiade
		2010.05-10	Expo 2010 Shanghai (World Expo)
		2010.11	2010 Asian Games
		2011.08	2011 Summer Universiade
		2013.10	2013 East Asian Games
		2014.08	2014 Summer Youth Olympics
Jevent	events in Japan	1998.02	Nagano 1998 Winter Olympics
		2001.05	2001 East Asian Games
		2001.08	2001 World Games
		2002.05-06	2002 FIFA World Cup
		2005.03-09	Expo 2005 AICHI (World Expo)

Kevent	events in Korea	1997.01	1997 Winter Universiade
		1997.05	1997 East Asian Games
		2002.05-06	2002 FIFA World Cup
		2002.09-10	2002 Asian Games
		2003.08	2003 Summer Universiade
		2012.05-08	Expo 2012 Yeosu (Specialised Expo)
		2014.09-10	2014 Asian Games
		2015.07	2015 Summer Universiade
		2015.10	2015 Military World Games
		2018.02	Pyeongchang 2018 Winter Olympics

2.3 Stationarity Tests

There are two types of *non-stationary* time series. Integrated (also known as **difference-stationary**) process, which is stationary after differencing. The other type is **trend-stationary** process, which is stationary after detrending or removing a deterministic trend. Please refer to Chapter 15 and Chapter 16 of the classic textbook on Time Series Analysis (James Hamilton, 1994, Princeton University Press) for details.

In our case, the (log) tourist flows are trend-stationary. "Spurious regressions" are only associated with difference-stationary (random walk) processes, so our regressions are not spurious. An eyeball test of Figure 3 show that both (log) tourist flows are stationary around a log-linear trend and the deviations from the trend are very short-lived and converging back to the trend. To make this point clearer, we report Augmented Dickey-Fuller (ADF) tests of the two tourist flows ($\ln CJ$ = log tourist flow to Japan; $\ln CK$ = log tourist flow to Korea) to formally confirm our argument. The null hypothesis of the ADF test is that "the process is integrated of order 1 after controlling for a determinist trend". The p-value is 0.0001, significantly rejecting the null. Therefore, the tourist flow is trend-stationary, NOT difference-stationary (integrated). The same applies to other dependent variables such as log GDP per capita, log exchange rates of each pair, and log oil price.

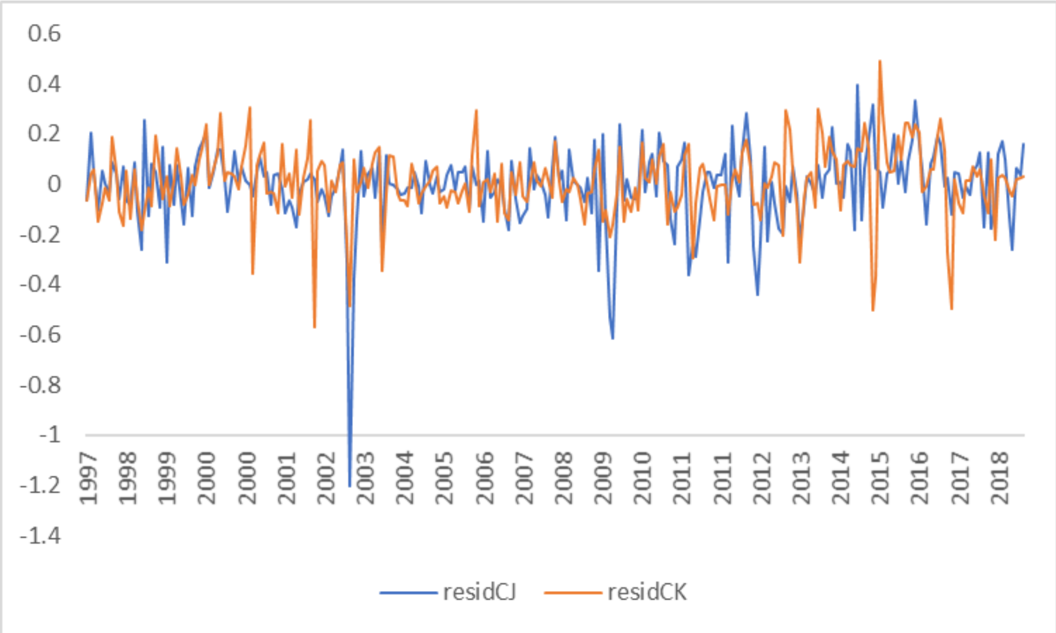
Variable	**Null Hypothesis**	**Test Statistics**	**MacKinnon P-value**
$\ln F_{CJ}$	Random walk with trend	-5.242	0.0001
$\ln F_{CK}$	Random walk with trend	-5.150	0.0001
$\ln GDPP_C$	Random walk with trend	-5.999	0.0000
$\ln GDPP_K$	Random walk with trend	-5.482	0.0000
$\ln GDPP_J$	Random walk with drift	-2.820	0.0026
$\ln EX_{CJ}$	Random walk with drift	-2.192	0.0146
$\ln EX_{CK}$	Random walk with drift	-3.527	0.0002
$\ln OIL$	Random walk with drift	-1.719	0.0434

2.4 Heteroscedasticity Tests

The assumption of GARCH can be tested formally. Note that the MGARCH model is a 2-equation system, so there are two residuals, one for the China-Japan equation and one for the China-Korea equation, respectively. We perform the Engel (1982) procedure to confirm the ARCH assumption and verify that the regression is not spurious by the stationarity of the residuals.

1) *Residuals from the original model.* As shown below, both residCJ (blue) and residCK (orange) tend to have volatility clustering, i.e. high volatility period tends to follow high volatility period (e.g. after 2010).

2) *Heteroskedasticity test.* We follow the Engel (1982) procedure to create squared residuals for residCJ2 and residCK2, and regress on their own lags. The null hypothesis is “there is no ARCH in the residuals”, which is equivalent to assuming a joint zero of all the regressors (lagged squared residuals). We report the test results for a lag length of 2, but it holds for a variety of lag lengths. Given that the null is significantly rejected (P-values of the joint tests are smaller than 5%), we can conclude that ARCH structure exists in both equations.



residCJ2	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
residCJ2						
L1.	.1751586	.0624849	2.80	0.005	.0521086	.2982086
L2.	-.017747	.0624866	-0.28	0.777	-.1408002	.1053062
_cons	.0205785	.0062751	3.28	0.001	.0082211	.0329359

```
. test L1.residCJ2 = L2.residCJ2 = 0

 ( 1)  L.residCJ2 - L2.residCJ2 = 0
 ( 2)  L.residCJ2 = 0

       F(  2,   256) =    3.95
            Prob > F =    0.0205
```

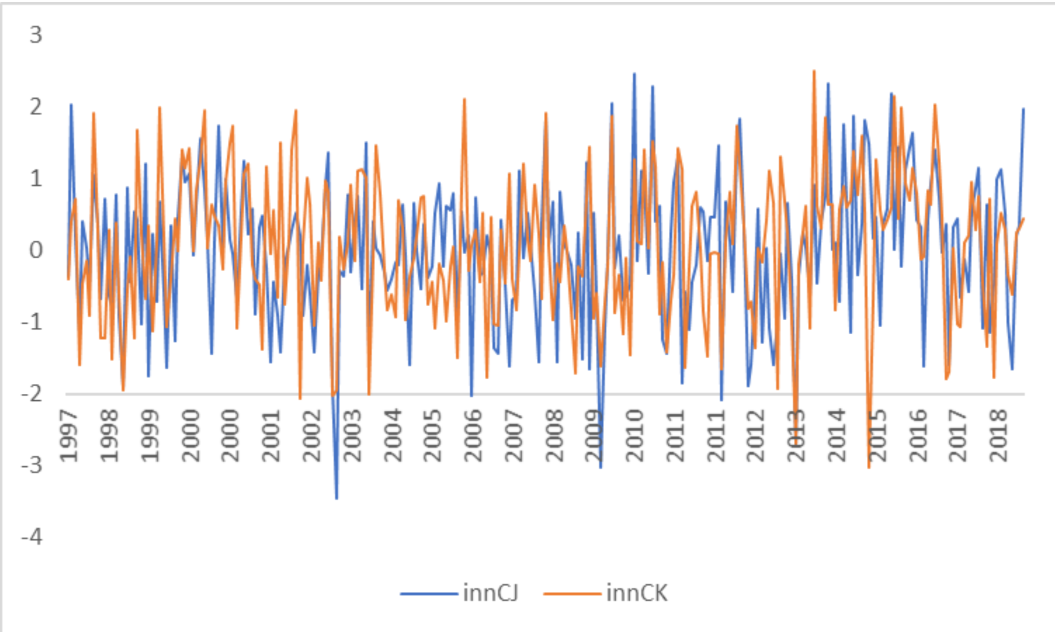
residCK2	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
residCK2						
L1.	.2662718	.0623909	4.27	0.000	.143407	.3891367
L2.	.0588087	.0623993	0.94	0.347	-.0640726	.18169
_cons	.0129701	.0028472	4.56	0.000	.0073631	.0185771

```
. test L1.residCK2 = L2.residCK2 = 0

 ( 1)  L.residCK2 - L2.residCK2 = 0
 ( 2)  L.residCK2 = 0

       F(  2,   256) =   11.62
            Prob > F =    0.0000
```

***3)** Residuals of the MGARCH model.* After controlling for the ARCH structure, we have obtained the innovations. Both innCJ (blue) and innCK (orange) fluctuate with a similar magnitude of volatility throughout the sample period and behaving like a white noise. To formally test that, we conduct the Engle (1982) test for both innovations. As shown in the two joint tests below, P-values are greater than 5% and the null hypothesis (no ARCH) cannot be rejected. We report the conclusion for a lag of 2 periods, but this conclusion holds for other lengths of lags.



innCJ2	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
innCJ2						
L1.	-.0484614	.062188	-0.78	0.437	-.1709266	.0740039
L2.	-.0885501	.0621876	-1.42	0.156	-.2110145	.0339144
_cons	1.127141	.1285714	8.77	0.000	.8739487	1.380333

```
. test L1.innCJ2 = L2.innCJ2 = 0

 ( 1)  L.innCJ2 - L2.innCJ2 = 0
 ( 2)  L.innCJ2 = 0

       F(  2,   256) =    1.27
            Prob > F =    0.2825
```

innCK2	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
innCK2						
L1.	-.0840308	.0624503	-1.35	0.180	-.2070125	.038951
L2.	-.0391267	.0624574	-0.63	0.532	-.1621225	.0838691
_cons	1.130402	.1241347	9.11	0.000	.8859464	1.374857

```
. test L1.innCK2 = L2.innCK2 = 0

 ( 1)  L.innCK2 - L2.innCK2 = 0
 ( 2)  L.innCK2 = 0

       F(  2,   256) =    1.04
            Prob > F =    0.3549
```

2.5 Lag Length

To provide statistical evidence for choosing a lag length of 1, we compare a variety of information criteria for VARs with different lag lengths below. It is shown that according to the two popular information criteria (SBIC and HQIC), a lag length of 1 is optimal, though a lag length of 2 is preferred in terms of AIC and FPE. For the sake of parsimony and the principle of Occam’s razor, we opt for a lag length of 1 in the VAR structure.

```
. varsoc lnCJ lnCK, exog($z L(0/3).(CJ5 CK5)) m(3)

 Selection-order criteria
 Sample:  1997m4 - 2018m12                    Number of obs      =       261
```

lag	LL	LR	df	p	FPE	AIC	HQIC	SBIC
0	55.5389				.004155	.187442	.626621	1.28001
1	314.865	518.65	4	0.000	.000588	-1.76908	-1.30795*	-.621881*
2	321.694	13.656*	4	0.008	.000576*	-1.79075*	-1.30766	-.588924
3	325.189	6.9917	4	0.136	.000578	-1.78689	-1.28184	-.530432

```
 Endogenous:  lnCJ lnCK
  Exogenous:  t m2 m3 m4 m5 m6 m7 m8 m9 m10 m11 m12 Cdisast L.Cdisast
              Jdisast L.Jdisast Kdisast L.Kdisast Cevent L.Cevent Jevent
              L.Jevent Kevent L.Kevent lnCGDP lnJGDP lnKGDP Choliday lneCJ
              lneCK lnoil CJ5 L.CJ5 L2.CJ5 L3.CJ5 CK5 L.CK5 L2.CK5 L3.CK5
              _cons
```

Most empirical studies adopting MGARCH assumption of the error distribution, including the original seminal papers by Bollerslev, Engle and Wooldridge (1988) and Bollerslev (1990), only use MGARCH(1, 1) in their empirical models. There are two reasons for this choice. First, MGARCH(1, 1) is able to capture most dynamic features in most medium-frequency data. Second, while higher order of lags may improve the goodness of fit, even a MGARCH(2, 2) would tremendously increase the number of parameters to be estimated and make the estimation convergence extremely difficult (MGARCH models are estimated by maximum likelihood algorithms numerically). It is not like simple VARMA models estimated by least squares analytically, in which case information criteria are usually used to choose the optimal lag lengths among a variety of combinations of lag lengths.

References

Anderson, J.; & Van Wincoop, E. (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review*, *93*(1), 170-192.

Bollerslev, T., Engle, R., Wooldridge, J. (1988) A capital asset pricing model with time-varying covariances. *Journal of Political Economy*, *96*(1), 116-131.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *The Review of Economics and Statistics*, *72*(3), 498-505.

McCallum, J. (1995). National borders matter: Canada-U.S. regional trade patterns. *American Economic Review*, *85*(3), 615-623.

Kimura, F., Lee, H. (2006). The gravity equation in international trade in services. *Review of World Economics*. *142*, 92-121.

Morley, C., Rosselló, J., & Santana-Gallego, M. (2014). Gravity models for tourism demand: Theory and use. *Annals of Tourism Research, 48,* 1-10.

Hoti, S., McAleer, M., & Shareef, R. (2007). Modelling international tourism and country risk spillovers for Cyprus and Malta. *Tourism Management, 28*(6), 1472-1484.

Sims, C. A. (1980). Macroeconomics and Reality. *Econometrica*, *48*(1), 1-48.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*(4), 987-1007.

Santamaria, D., & Filis, G. (2019). Tourism demand and economic growth in Spain: New insights based on the yield curve. *Tourism Management, 75,* 447-459.

Appendix: Mathematical Derivation

This Appendix elaborates the mathematical details of the structural model.

- **Derive the Theoretical Model**

Assuming a Cobb-Douglas utility, the maximization problem is:

$$\max_{f_J(i), f_K(i)} f_J(i)^{\alpha_J} f_K(i)^{\alpha_K}, \text{ subject to: } c_J f_J(i) + c_K f_K(i) = m(i)^{\mu}$$

Set up the Lagrangian:

$$L \equiv f_J(i)^{\alpha_J} f_K(i)^{\alpha_K} - \lambda\left[c_J f_J(i) + c_K f_K(i) - m(i)^{\mu}\right]$$

Take partial derivative with respect to $f_J(i)$, $f_K(i)$ and λ, resulting in the optimality conditions:

$$\frac{\partial L}{\partial f_J(i)} = \alpha_J f_J(i)^{\alpha_J - 1} f_K(i)^{\alpha_K} - \lambda c_J = 0 \quad (13)$$

$$\frac{\partial L}{\partial f_K(i)} = \alpha_K f_J(i)^{\alpha_J} f_K(i)^{\alpha_K - 1} - \lambda c_K = 0 \quad (14)$$

$$\frac{\partial L}{\partial \lambda} = c_J f_J(i) + c_K f_K(i) - m(i)^{\mu} = 0 \quad (15)$$

Combine the first two optimality conditions (13) and (14) to substitute out the Lagrangian multiplier λ, and then combine with (15) to solve for the two individual optimal flows $f_d(i)^*$, $d = J, K$. Aggregate the individual flows over i to get the country-level flows F_d.

$$f_d(i)^* = \alpha_d \frac{m(i)^{\mu}}{c_d} \rightarrow F_d = \sum_i f_d(i)^* = \alpha_d \frac{\sum_i m(i)^{\mu}}{c_d}$$

Substitute in the definition of α_d and the specifications for A_d and c_d, we have derived the aggregate tourist flow to destination d:

$$F_d = \frac{A_d}{A_J + A_K} \frac{m(i)^{\mu}}{c_d} = \frac{A_d(GDPP_d, E_d, D_d, PR_d)}{A_J(GDPP_J, E_J, D_J, PR_J) + A_K(GDPP_K, E_K, D_K, PR_K)} \frac{\sum_i m(i)^{\mu}}{c_d(z_d, PR_d)} \quad (16)$$

- **Properties of the Optimal Tourist Flows**

Before proceeding with (16) further for an empirical model, we digress here to derive the two properties (equation (10) and (11) in the main text). First, differentiate the China-Japan flow (F_J) of solution (16) with respect to PR_J to obtain the *bilateral* political effect:

$$F_J = \frac{A_J}{A_J+A_K}\frac{m(i)^\mu}{c_J} \rightarrow \frac{\partial F_J}{\partial PR_J} = \frac{\partial F_J}{\partial A_J} \times \frac{\partial A_J}{\partial PR_J} + \frac{\partial F_J}{\partial c_J} \times \frac{\partial c_J}{\partial PR_J}$$

Now work out each component on the right-hand side:

$$\frac{\partial F_J}{\partial A_J} = \left[\frac{1}{A_J+A_K} - \frac{A_J}{\left(A_J+A_K\right)^2}\right]\frac{m(i)^\mu}{c_J} = \frac{A_K}{A_J\left(A_J+A_K\right)}\frac{A_J}{A_J+A_K}\frac{m(i)^\mu}{c_J} = \frac{A_K F_J}{A_J\left(A_J+A_K\right)} > 0$$

$$\frac{\partial A_J}{\partial PR_J} > 0$$

$$\frac{\partial F_J}{\partial c_J} = -\frac{A_J}{A_J+A_K}\frac{m(i)^\mu}{c_J^2} = -\frac{1}{c_J}\frac{A_J}{A_J+A_K}\frac{m(i)^\mu}{c_J} = -\frac{F_J}{c_J} < 0$$

$$\frac{\partial c_J}{\partial PR_J} < 0$$

Therefore, we have proved:

$$\frac{\partial F_J}{\partial PR_J} = \frac{\partial F_J}{\partial A_J} \times \frac{\partial A_J}{\partial PR_J} + \frac{\partial F_J}{\partial c_J} \times \frac{\partial c_J}{\partial PR_J} = \underbrace{\frac{A_K F_J}{A_J\left(A_J+A_K\right)}}_{>0} \times \underbrace{\frac{\partial A_J}{\partial PR_J}}_{>0} + \underbrace{\left(-\frac{F_J}{c_J}\right)}_{<0} \times \underbrace{\frac{\partial c_J}{\partial PR_J}}_{<0} > 0$$

Similarly, differentiate F_K of (16) with respect to PR_J to obtain the *multilateral* political effect:

$$\frac{\partial F_K}{\partial A_J} = -\frac{A_K}{\left(A_J+A_K\right)^2}\frac{m(i)^\mu}{c_K} = -\frac{1}{A_J+A_K}\frac{A_K}{A_J+A_K}\frac{m(i)^\mu}{c_K} = -\frac{F_K}{A_J+A_K} < 0$$

$$\frac{\partial A_J}{\partial PR_J} > 0$$

$$\rightarrow \frac{\partial F_K}{\partial PR_J} = \frac{\partial F_K}{\partial A_K} \times \frac{\partial A_K}{\partial PR_K} = \underbrace{\left(-\frac{F_K}{A_J+A_K}\right)}_{<0} \times \underbrace{\frac{\partial A_K}{\partial PR_K}}_{>0} < 0$$

- **Derive the Empirical Model**

To advance (16) further for the empirical analysis, we need to make use of two approximations. First, let's focus on the term $\frac{A_J}{A_J+A_K}$, which can be log-linearized in the neighborhood of $\frac{A_K}{A_J}=1$. We use 1 as the center of approximation because Japan and Korea belong to the same tourism category, so the ratio of attractiveness must be close to 1 in the steady state.

$$\ln\frac{A_J}{A_J+A_K}=\ln\frac{1}{1+\frac{A_K}{A_J}}=-\ln\left(1+\frac{A_K}{A_J}\right)\approx constant-\ln\frac{A_K}{A_J} \quad \text{(17)a}$$

The last step in (17)a is based on the observation that the difference between $-\ln\left(1+\frac{A_K}{A_J}\right)$ and $\ln\frac{A_K}{A_J}$ ranges between 0.6 and 0.8 if $\frac{A_K}{A_J}$ only deviates from 1 by $\pm 20\%$. Relative to the magnitude of $\ln F$ (the mean is around 11, see **Table 2**), the magnitude of the variation is ignorable, $\frac{0.8-0.6}{11}=1.8\%$. Thus, we can approximately treat $\ln\frac{A_K}{A_J}-\ln\left(1+\frac{A_K}{A_J}\right)$ as a constant in (17)a.

The second approximation is about $\sum_i m(i)^\mu$. In principle, the sum of a function is not equal to the function of a sum according to the Jensen's inequality: $\sum_i m(i)^\mu \neq (\sum_i m(i))^\mu$. To make use of the fact that the sum of individual income of the origin country ($\sum_i m(i)$) is equal to GDP of the origin country (China), we need to compensate for the gap by a multiplicative fixed proportion: $\sum_i m(i)^\mu=\bar{M}\times(\sum_i m(i))^\mu$. Denote GDP per capita by $GDPP_C$, population by POP_C, and take logarithms: $\ln\sum_i m(i)^\mu=\ln\bar{M}+\mu\ln GDPP_C+\mu\ln POP$. Note that POP_C grows exponentially at a stable rate during our sample period, so $\ln POP_C$ can be merged into a deterministic term, which has a constant component, a trend component and a seasonality component. Approximately, we have:

$$\ln\sum_i m(i)^\mu\approx\underbrace{constant+trend+seasonality}_{deterministic}+\mu\ln GDPP_C \quad \text{(17)b}$$

Make use of (17)a and (17)b to log-linearize (16), we have:

$$\ln F_J\approx\boldsymbol{\beta0}_J'\mathbf{x0}_t+\mu\ln GDPP_C+\ln A_J-\ln A_K-\ln c_J \quad \text{(18)J}$$

$$\ln F_K\approx\boldsymbol{\beta0}_K'\mathbf{x0}_t+\mu\ln GDPP_C-\ln A_J+\ln A_K-\ln c_K \quad \text{(18)K}$$

All the constant, trend and seasonality terms of the two equations are collected in the deterministic term $\mathbf{x0}_t$, with $\boldsymbol{\beta0}_J$ and $\boldsymbol{\beta0}_K$ being the coefficients.

Lastly, substitute the assumed power-exponential specifications (19)JK and (20)JK into (18)JK:

$$A_J = A\left(GDPP_J, E_J, D_J, PR_J\right) \equiv GDPP_J^{\gamma_1} \exp\left(\gamma_2 E_J\right) \exp\left(\gamma_3 D_J\right) \exp\left(\gamma_4 PR_J\right) \qquad \text{(19)J}$$

$$A_K = A(GDPP_K, E_K, D_K, PR_K) \equiv GDPP_K^{\delta_1} \exp(\delta_2 E_K) \exp(\delta_3 D_K) \exp(\delta_4 PR_K) \qquad \text{(19)K}$$

$$c_J = c\left(\mathbf{z}_J, PR_J\right) \equiv \exp\left(\boldsymbol{\gamma}_5' \ln \mathbf{z}_J\right) \exp\left(-\gamma_6 PR_J\right) \qquad \text{(20)J}$$

$$c_K = c(\mathbf{z}_K, PR_K) \equiv \exp(\boldsymbol{\delta}_5' \ln \mathbf{z}_K) \exp(-\delta_6 PR_K) \qquad \text{(20)K}$$

We have then derived the multilateral gravity equation:

$$\begin{gathered}\ln F_J \approx \boldsymbol{\beta 0}_J' \mathbf{x0}_t + \mu_J \ln GDPP_C + \gamma_1 \ln GDPP_J + \gamma_2 E_J + \gamma_3 D_J - \\ \delta_1 \ln GDPP_K - \delta_2 E_K - \delta_3 D_K + (\gamma_4 + \gamma_6) PR_J - \delta_4 PR_K - \boldsymbol{\gamma_5'} \ln \mathbf{z}_J\end{gathered} \qquad \text{(21)J}$$

$$\begin{gathered}\ln F_K \approx \boldsymbol{\beta 0}_K' \mathbf{x0}_t + \mu_K \ln GDPP_C - \gamma_1 \ln GDPP_J - \gamma_2 E_J - \gamma_3 D_J + \\ \delta_1 \ln GDPP_K + \delta_2 E_K + \delta_3 D_K - \gamma_4 PR_J + (\delta_4 + \delta_6) PR_K - \boldsymbol{\delta}_5' \ln \mathbf{z}_K\end{gathered} \qquad \text{(21)K}$$

After redefining coefficients and variables, (21)JK can be rewritten in the matrix form:

$$\underbrace{\begin{bmatrix}\ln F_J \\ \ln F_K\end{bmatrix}}_{\equiv \mathbf{y}_t} = \underbrace{\begin{bmatrix}\boldsymbol{\beta 0}_J' & \boldsymbol{\beta 1}_J' & \boldsymbol{\beta 2}_J' \\ \boldsymbol{\beta 0}_K' & \boldsymbol{\beta 1}_K' & \boldsymbol{\beta 2}_K'\end{bmatrix}}_{\equiv \mathbf{B}} \underbrace{\begin{bmatrix}\mathbf{x0}_t \\ \mathbf{x1}_t \\ \mathbf{x2}_t\end{bmatrix}}_{\equiv \mathbf{x}_t} + \underbrace{\begin{bmatrix}\epsilon_J \\ \epsilon_K\end{bmatrix}}_{\equiv \boldsymbol{\epsilon}_t}, \text{ or just } \mathbf{y}_t = \mathbf{B}\mathbf{x}_t + \boldsymbol{\epsilon}_t. \qquad (22)$$

$\mathbf{x0}_t$ is a vector of deterministic terms including a constant, a trend and seasonality dummies; $\mathbf{x1}_t$ is a vector of *unilateral* covariates including GDP per capita (GDPP), events/disasters in each country and holidays in China (time costs of travel); $\mathbf{x2}_t$ is a vector of *bilateral* covariates including political relation between each country pair, exchange rates and oil price (monetary costs of travel). An error term $\boldsymbol{\epsilon}_t$ is added to the equation to capture all residuals that cannot be explained by these covariates. The coefficient matrix **B** corresponds to the coefficients in (21)JK, e.g. $\boldsymbol{\beta 1}_J = \left[\mu_J, \gamma_1, \gamma_2, \gamma_3, -\delta_1, -\delta_2, -\delta_3\right]$ and $\boldsymbol{\beta 2}_J = [\gamma_4 + \gamma_6, -\delta_4, \boldsymbol{\gamma}_5]$.