Appendix to "Gravity, trade integration, and heterogeneity across industries", Natalie Chen and Dennis Novy, *Journal of International Economics*.

Appendix A

Table A1: The Determinants of EU Trade Integration: Robustness

	(1)	(2)	(3)	(4)	(5)
Elasticities	σ_k	σ_k	σ_k	σ_k	σ_k
Dependent variable	$\ln \theta^k_{ij,t}$	$\ln \theta^k_{ij,t}$	$\ln \theta^k_{ij,t}$	$\ln \theta^k_{ij,t}$	$\ln \theta^k_{ij,t}$
Geography/Transport Costs					
$\ln D_{ij}$	–	0.378^{a} (11.305)	0.352^{a} (11.715)	0.382^{a} (11.288)	0.381^{a} (11.265)
$\ln (D_{ii} \times D_{jj})$	–	-0.926^{a} (−6.901)	-0.845^{a} (−7.134)	-0.928^{a} (−6.910)	-0.929^{a} (−6.815)
Adj_{ij}	–	-0.096^{a} (−2.910)	-0.089^{a} (−2.998)	-0.093^{a} (−2.759)	-0.094^{a} (−2.821)
$Lang_{ij}$	–	-0.325^{a} (−5.667)	-0.314^{a} (−5.975)	-0.326^{a} (−5.691)	-0.328^{a} (−5.655)
$\ln wv^k_t$	0.421^{a} (3.961)	0.386^{a} (3.577)	0.407^{a} (4.212)	0.385^{a} (3.571)	0.414^{a} (3.387)
$\ln cfob^k_t$	0.023^{a} (2.930)	0.018^{b} (2.406)	0.015^{b} (2.218)	0.019^{b} (2.451)	0.024^{b} (2.070)
$Island_{ij}$	–	–	–	0.004 (0.113)	–
Policy Variables					
FI, AT_{ij}	–	0.229^{a} (8.698)	0.217^{a} (9.151)	0.230^{a} (8.690)	0.231^{a} (8.663)
$noEURO_{ij,t}$	–	0.005 (0.265)	0.008 (0.451)	0.007 (0.306)	0.007 (0.336)
$Schengen_{ij,t}$	–	-0.128^{a} (−3.353)	-0.107^{a} (−3.162)	-0.124^{a} (−3.179)	-0.133^{a} (−3.395)
$\ln TBT^k_{ij}$	0.335^{b} (2.202)	0.386^{a} (3.071)	0.339^{a} (3.031)	0.382^{a} (3.032)	0.374^{a} (2.984)
$\ln Proc^k_{ij,t}$	-4.184^{a} (−3.208)	-3.836^{a} (−3.202)	-3.387^{a} (−3.285)	-3.869^{a} (−3.225)	-4.825^{a} (−3.317)
$\ln VAT^k_{ij}$	-21.369^{a} (−6.601)	-13.047^{a} (−5.891)	-10.906^{a} (−5.440)	-13.192^{a} (−5.655)	-13.372^{a} (−5.891)
Other Costs					
$\ln Prod^k_{ij,t-1}$	-0.494^{a} (−7.875)	-0.354^{a} (−6.984)	-0.347^{a} (−7.516)	-0.353^{a} (−6.967)	-0.363^{a} (−6.952)
$Zeros^k_{ij,t}$	2.664^{a} (31.371)	–	2.859^{a} (15.444)	2.789^{a} (16.667)	2.788^{a} (16.389)
Controls					
$KS^k_{ij,t}$	0.042 (0.834)	0.027 (0.579)	0.023 (0.549)	0.026 (0.554)	0.028 (0.581)
$\ln Goods^k$	-0.331^{a} (−14.071)	-0.337^{a} (−14.227)	-0.320^{a} (−14.887)	-0.340^{a} (−14.320)	-0.338^{a} (−14.099)
Estimator	OLS	OLS	OLS	OLS	OLS
Fixed Effects	$t \times ij, K$	t, K	t, K	t, K	$t \times K$
Weighted	No	No	Yes	No	No
Sample	Full	Excl. zeros	Full	Full	Full
N	12,116	12,104	12,116	12,116	12,116
Adj-R^2	0.732	0.722	0.718	0.722	0.723

(continued on next page)

Table A1: The Determinants of EU Trade Integration: Robustness (continued)

	(6)	(7)	(8)	(9)	(10)	(11)
Elasticities	σ_k	$\sigma_k + se$	$\sigma_k - se$	σ_k^{IM}	σ_k^{BW}	σ_k
Dependent variable	$\ln \theta_{ij,t}^k$	$\ln \theta_{ij,t}^k$	$\ln \theta_{ij,t}^k$	$\ln \theta_{ij,t}^k$	$\ln \theta_{ij,t}^k$	$\ln \theta_{ij,t}^k$
Geography/Transport Costs						
$\ln D_{ij}$	0.380^a (9.792)	0.341^a (10.857)	0.455^a (11.256)	0.490^a (20.581)	0.997^a (9.294)	0.306^a (7.610)
$\ln (D_{ii} \times D_{jj})$	-1.064^a (-7.108)	-0.772^a (-6.126)	-1.257^a (-7.604)	-1.243^a (-15.147)	-1.816^a (-3.389)	-0.823^a (-5.925)
Adj_{ij}	-0.086^b (-2.323)	-0.095^a (-3.033)	-0.081^b (-2.129)	-0.049^b (-2.194)	-0.053 (-0.584)	-0.144^a (-3.964)
$Lang_{ij}$	-0.341^a (-5.627)	-0.286^a (-5.230)	-0.395^a (-6.106)	-0.393^a (-9.846)	-0.499^b (-2.477)	-0.335^a (-5.764)
$\ln wv_t^k$	0.393^a (3.580)	0.337^a (3.131)	0.459^a (4.209)	0.431^a (13.184)	-1.202^a (-2.888)	0.501^a (4.748)
$\ln cfob_t^k$	0.024^b (2.542)	0.016^b (2.187)	0.024^a (2.938)	0.012^b (2.388)	0.151^a (4.199)	0.019^b (2.422)
$Island_{ij}$	–	–	–	–	–	–
Policy Variables						
FI, AT_{ij}	0.255^a (8.287)	0.210^a (8.483)	0.257^a (8.196)	0.264^a (14.118)	0.208^a (2.613)	0.281^a (9.550)
$noEURO_{ij,t}$	0.006 (0.267)	0.006 (0.324)	0.011 (0.445)	0.032^b (2.243)	-0.177^a (-2.594)	0.023 (1.185)
$Schengen_{ij,t}$	-0.107^b (-2.429)	-0.131^a (-3.625)	-0.118^a (-2.611)	-0.124^a (-4.431)	-0.272^b (-1.967)	-0.153^a (-3.936)
$\ln TBT_{ij}^k$	0.392^a (2.865)	0.434^a (3.411)	0.301^b (2.387)	0.213^a (8.590)	0.611^b (1.984)	0.368^a (2.771)
$\ln Proc_{ij,t}^k$	-4.772^b (-2.559)	-3.460^a (-3.071)	-4.708^a (-3.523)	-2.106^c (-1.895)	-0.282 (-0.093)	-4.266^a (-3.386)
$\ln VAT_{ij}^k$	-14.207^a (-5.744)	-9.634^a (-4.940)	-22.175^a (-6.927)	-7.068^a (-5.869)	8.406^c (1.701)	-13.766^a (-6.162)
Other Costs						
$\ln Prod_{ij,t-1}^k$	-0.379^a (-6.956)	-0.298^a (-6.320)	-0.424^a (-7.113)	-0.272^a (-8.792)	0.028 (0.250)	-0.770^{aa} (-5.178)
$Zeros_{ij,t}^k$	2.793^a (16.855)	2.035^a (44.092)	4.378^a (5.372)	2.225^a (12.208)	3.764^a (7.621)	2.820^a (17.644)
Controls						
$KS_{ij,t}^k$	0.009 (0.178)	0.037 (0.853)	-0.007 (-0.120)	0.018 (0.548)	0.183 (1.367)	0.135^b (2.341)
$\ln Goods^k$	-0.338^a (-14.263)	-0.280^a (-12.744)	-0.471^a (-15.270)	-0.195^a (-18.753)	-0.541^a (-6.380)	-0.335^a (-14.405)
Estimator	OLS	OLS	OLS	OLS	OLS	IV
Fixed Effects	t, K	t, K	t, K	$t, \widetilde{K}$	t, K	t, K
Weighted	No	No	No	No	No	No
Sample	Excl. $\phi_{ij,t}^k$	Full	Full	Full	Full	Full
N	9,576	12,116	12,116	11,924	12,116	12,060
Adj-R^2	0.596	0.724	0.699	0.644	0.550	0.715

Notes: In all columns the dependent variable is $\ln \theta_{ij,t}^k$. Weighted regression in (3) where the weights are the product of sectoral bilateral exports in the initial year 1999. In (6) the sample "Excl. $\phi_{ij,t}^k$" means that we exclude from the sample the 37 industries for which the variation in $\theta_{ij,t}^k$ to be explained is the same as the variation in *phi*-ness once 3-digit fixed effects are included. In (7) and (8) the elasticities of substitution used to compute $\theta_{ij,t}^k$ are equal to σ_k plus or minus one standard error, respectively. In (9) the elasticities of substitution σ_k^{IM} used to compute $\theta_{ij,t}^k$ are from Imbs and Méjean (2009) while in (10) the elasticities σ_k^{BW} are from Broda and Weinstein (2006). In (11) labor productivity $\ln Prod_{ij,t-1}^k$ is instrumented by average firm size, R&D and sectoral productivity for the US. The fixed effects t, ij, K and $\widetilde{K}$ denote year, country pair, 3-digit and 2-digit fixed effects, respectively. Robust standard errors are adjusted for clustering at the 4-digit NACE Rev.1 level in each country pair. The sample period is 1999-2003. t-statistics in parentheses. Constant terms are included but not reported. ^{a}, ^{b} and ^{c} indicate significance at 1, 5 and 10 percent levels, respectively.

Table A2: The Determinants of EU Trade Integration: Cross-Sectional Samples

	(1)	(2)	(3)	(4)	(5)
Year	1999	2000	2001	2002	2003
Dependent variable	$\ln \theta^k_{ij,t}$	$\ln \theta^k_{ij,t}$	$\ln \theta^k_{ij,t}$	$\ln \theta^k_{ij,t}$	$\ln \theta^k_{ij,t}$
Geography/Transport Costs					
$\ln D_{ij}$	$0.455^{\ a}$ (12.996)	$0.388^{\ a}$ (10.521)	$0.388^{\ a}$ (11.010)	$0.366^{\ a}$ (10.762)	$0.381^{\ a}$ (11.274)
$\ln (D_{ii} \times D_{jj})$	-1.005^a (-6.948)	-1.010^a (-7.013)	-0.950^a (-6.693)	-0.906^a (-6.563)	-0.799^a (-5.817)
Adj_{ij}	-0.043 (-1.253)	-0.077^b (-2.278)	-0.099^a (-2.813)	-0.099^a (-2.912)	-0.090^a (-2.634)
$Lang_{ij}$	-0.309^a (-4.871)	-0.318^a (-5.507)	-0.330^a (-5.553)	-0.341^a (-5.782)	-0.342^a (-5.490)
$\ln wv^k_t$	0.416^a (3.075)	0.340^b (2.475)	0.497^a (3.996)	0.375^a (2.636)	0.389^a (2.882)
$\ln cfob^k_t$	0.054^a (2.933)	0.041^b (2.561)	0.021 (1.196)	0.072^a (3.442)	-0.018 (-1.270)
Policy Variables					
FI, AT_{ij}	0.168^a (6.067)	0.211^a (7.558)	0.236^a (8.491)	0.246^a (8.895)	0.234^a (8.386)
$noEURO_{ij,t}$	–	–	–	0.003 (0.102)	0.007 (0.195)
$Schengen_{ij,t}$	-0.114^a (-2.669)	-0.136^a (-3.254)	-0.145^a (-2.707)	-0.149^b (-2.268)	-0.181^a (-2.749)
$\ln TBT^k_{ij}$	0.312^a (2.906)	0.328^a (3.031)	0.364^a (2.754)	0.356^a (2.649)	0.457^a (3.029)
$\ln Proc^k_{ij,t}$	-9.517^b (-2.367)	-5.006^a (-2.622)	-4.512^b (-2.087)	-5.089^b (-2.268)	-4.911^a (-3.811)
$\ln VAT^k_{ij}$	-14.669^a (-6.082)	-16.745^a (-6.630)	-13.164^a (-5.774)	-11.479^a (-4.648)	-10.422^a (-4.140)
Other Costs					
$\ln Prod^k_{ij,t-1}$	–	-0.304^a (-4.826)	-0.393^a (-7.321)	-0.421^a (-7.267)	-0.361^a (-6.417)
$Zeros^k_{ij,t}$	–	–	–	$2.479^{\ a}$ (53.025)	$2.923^{\ a}$ (14.963)
Controls					
$KS^k_{ij,t}$	-0.080 (-1.256)	-0.078 (-1.423)	0.049 (0.802)	0.077 (1.231)	0.093 (1.420)
$\ln Goods^k$	$-0.330^{\ a}$ (-13.618)	$-0.338^{\ a}$ (-13.847)	$-0.346^{\ a}$ (-13.673)	$-0.314^{\ a}$ (-12.897)	$-0.345^{\ a}$ (-13.678)
N	3,029	3,029	3,029	3,029	3,029
Adj-R^2	0.719	0.722	0.722	0.727	0.724

Notes: The dependent variable is $\ln \theta^k_{ij,t}$. 3-digit industry fixed effects are included in all regressions. The sample period is 1999-2003. t-statistics in parentheses. Constant terms are included but not reported. ^{a}, ^{b} and ^{c} indicate significance at 1, 5 and 10 percent levels, respectively.

Appendix B: Data Appendix

Trade data Bilateral and total export and import trade flows (thousand Euros), and their corresponding weight (tons) are used for 11 EU countries between 1999 and 2003. The data are collected at the 8-digit Combined Nomenclature (CN) level of products which is the goods classification used within the EU for the purposes of foreign trade statistics. It was introduced in 1988 and is based on the Harmonized System (HS) international classification (there is full agreement between the two classifications at the 6-digit level). The CN is an annually revised classification, and in 2003 it covered a total of 10,404 products which were hierarchically classified into 21 sections and 99 chapters. Our analysis is however performed at the 4-digit NACE Rev.1 level of manufacturing industries, so the correspondence between the NACE Rev.1 and the yearly CN classifications is used to aggregate the data from the CN to the NACE Rev.1 level (see http://ec.europa.eu/eurostat/ramon). *Source: Eurostat COMEXT database.*

Weight-to-value and *c.i.f./f.o.b.* The bilateral weight-to-value ratio of exports (kilograms per Euro exported) is calculated separately for the two partners in each NACE Rev.1 industry and in each year, and is then averaged across all partners. We calculate the log of one plus weight-to-value as weight-to-value often takes on values much smaller than one.

The ratio between bilateral import ("Costs, Insurance and Freight," *c.i.f.*) and export ("Free On Board," *f.o.b.*) flows is calculated separately for the two partners in each NACE Rev.1 industry and year, and is then averaged across all partners, dropping the few cases where the ratio is smaller than one. We then use the log of the ratio. *Source: Eurostat COMEXT database.*

Elasticities of Substitution To estimate the elasticities of substitution, bilateral *c.i.f.* import trade flows (thousand Euros) and their corresponding weight (tons) are used for 11 EU countries importing from 46 countries between 1999 and 2003 at the 8-digit Combined Nomenclature (CN) level of manufacturing products. The 46 exporting countries include the 11 importing EU countries plus Argentina, Australia, Belgium-Luxembourg, Bulgaria, Canada, Chile, China, Colombia, Cyprus, Greece, Hong Kong, Hungary, Iceland, India, Indonesia, Israel, Japan, Malaysia, Malta, Mexico, New Zealand, North Korea, Norway, Pakistan, the Philippines, Poland, Romania, Saudi Arabia, Singapore, Sweden, Switzerland, Taiwan, Thailand, Turkey and the US. The sample includes 4,272,279 observations. *Source: Eurostat COMEXT database.*

Output Gross value of output (million Euros) at the 4-digit NACE Rev.1 level. *Source: Eurostat New Cronos database.*

Gravity Variables Dummies for sharing a common land border and a common (official) language. *Source: Centre d'Etudes Prospectives et d'Informations Internationales (CEPII).* International and domestic distances are weighted averages of the distances between regions using GDP shares as weights. *Source: Chen (2004).*

Schengen Agreement For each country, a dummy is equal to one in the years in which the provisions of the Schengen Agreement are fully implemented. We then take the average between the two partners. The resulting variable takes on values of zero, one half and one. The years of first implementation are chosen depending on the month in which the agreement went into force: Austria (1998), Denmark (2001), Finland (2001), France (1995), Germany (1995), Italy (1998), the Netherlands (1995), Portugal (1995), Spain (1995) while Ireland and the United Kingdom have not yet fully implemented the agreement.

Technical Barriers to Trade We use two sources. The European Commission's Eurobarometer reports opinions and experiences of European managers about the Single Market. A total of 4,900 managers at companies were interviewed by telephone in early 2006, the sample of companies being selected according to the size of countries and of companies, and the industry of activity. We use the answer to the question: "Could you tell me

whether you consider that for your company it is very important, rather important, rather unimportant or not important at all that future Single Market Policy tackles the question of *removing remaining technical barriers to trade in goods?*" For each country, we group the answers from all managers who replied that TBTs are indeed an important issue, and use the percentage so obtained as a country-specific indication on the relevance of TBTs. *Source: European Commission, 2006. Internal market: Opinions and experiences of businesses in EU-15. Flash Eurobarometer 180, TNS Sofres/EOS Gallup Europe.*

To capture the sectoral relevance of TBTs, industries are classified at the NACE70 level on a five-point scale according to the effectiveness of different measures undertaken by the Single Market Programme to eliminate TBTs: measures are successful and all significant barriers are removed (value of 1), measures are implemented and function well but some barriers remain (value of 2), measures are adopted but with implementation or transitional problems still to be overcome (value of 3), measures are proposed or implemented but not effective or with operating problems (value of 4), and no solution has been adopted (value of 5). Some industries are also identified as being not affected by TBTs prior to European integration (and are given a value of 1). Our industry-specific qualitative variable takes on values between 1 and 5, with larger values indicating a lack of market integration due to persisting TBTs. *Source: European Commission, 1998. Technical barriers to trade. The Single Market Review, Subseries III: Dismantling of Barriers 1.*

We then interact the log of one plus the share of managers replying that TBTs are important in country i with the log of one plus the share of managers replying that TBTs are important in country j and the industry-specific variable on TBTs.

Regarding the raw data used to compute the TBT variable, the survey of EU managers indicates that concerns about TBTs are the lowest in the UK (51 percent of interviewed managers consider TBTs as an important trade barrier), followed by Germany (59 percent), Finland (60 percent), France (61 percent), Austria (63 percent), Denmark (64 percent), the Netherlands (70 percent), Spain (71 percent), Italy (77 percent) and finally Portugal and Ireland (80 percent). Large and core EU countries such as the UK, Germany and France therefore tend to be less worried on average while periphery countries such as Portugal, Ireland and Spain are more concerned, Italy being an outlier to this classification. At the sector level, examples of industries where TBTs are successfully removed ($TBT^k = 1$) are "Dressing and dyeing of fur," "Electric domestic appliances," "Motor vehicles" or "Aircraft and spacecraft" while TBTs are still prevalent ($TBT^k = 5$) in "Jewellery" or "Imitation jewellery," among others.

Given that TBTs require specific product characteristics or production processes we would expect them to be stronger for differentiated than for homogeneous goods. The correlation between the sectoral indicator for TBTs, TBT^k, and the elasticity of substitution σ_k is indeed equal to -21 percent. A similar result arises when using the Rauch (1999) classification of products into three categories – commodities, reference priced and differentiated goods. The mean value for TBT^k is equal to 2.01 for differentiated products, which is significantly larger than the corresponding mean values of 1.21 and 1.24 for reference priced and homogeneous goods, respectively (the two means are not significantly different from each other).

Public Procurement We use two sources. For each country, we use the share of public procurement contracts advertized each year in the *Official Journal* of the EU. *Source: Eurostat.* The industries strongly affected by public procurement in Europe are identified at the NACE70 level and identified by a dummy variable. *Source: Davies and Lyons (1996).* We interact the log of one plus the proportion of public procurement contracts advertized in the *Official Journal* of the EU in country i with the log of one plus the proportion of public procurement contracts advertized in country j and the sectoral dummy.

Value-Added Taxes For each country, we use the standard VAT rate and replace it by the reduced VAT rates that apply to some categories of goods, as of January 1^{st}, 2008. We do not have information on the evolution of reduced VAT rates across goods over time, but we do not expect those to have changed much as the standard VAT rates changed very little during the time period we consider. We then compute the log of one plus the VAT rate that applies in each industry in country i with the log of one plus the VAT rate that

applies in each industry in country j. *Source: European Commission, 2008. VAT Rates Applied in the Member States of the European Community, DOC/2412/2008-EN.*

Productivity Real labor productivity is value-added divided by the number of employees at the 4-digit NACE Rev.1 level, deflated by GDP deflators. We then use the log of the average across partners of real labor productivity. *Source: Eurostat New Cronos database.*

Capital Shares Capital shares are value-added minus personnel costs, divided by personnel costs. We then calculate the log of one plus the absolute difference in capital shares between countries for each industry and in each year. *Source: Eurostat New Cronos database.*

Number of Product Categories *Source: Eurostat Prodcoms.*

GDP Deflators Value of GDP (million Euros) and volume of GDP (millions of 1995 constant Euros). The deflator is given by the ratio between the value and the volume of GDP. *Source: Eurostat New Cronos database.*

Instruments for Productivity First, we use US real labor productivity defined as value-added (deflated by the US CPI) divided by employment at the 4-digit 1987 SIC level. *Sources: NBER-CES Manufacturing Industry Database and OECD Main Economic Indicators.*

Second, we use average firm size given by the number of employees divided by the number of establishments at the 4-digit ISIC Rev.3 level. Missing observations for two industries decrease the sample by 70 observations. *Source: UNIDO Industrial Statistics Database.*

Third, we use the log of R&D expenditure as a share of GDP for country i times the log of R&D expenditure as a share of GDP for country j times the log of US shares of R&D spending in sectoral value-added. Nominal R&D expenditure and value-added for the US are available at the 2-digit ISIC Rev.3 level. *Source: OECD STAN database.* Gross domestic expenditure on R&D for EU countries is in millions of national currency. *Source: OECD Science and Technology Indicators.* The value of GDP for EU countries is in millions of national currency. *Source: Eurostat New Cronos database.*