

When Economies Fall Together

What 54 years of GDP growth in twenty OECD countries show about clustered contractions

Tal Zisckind

October 2026

For risk modelers, stress-test designers and fiscal-policy staff.

Abstract

Risk models often treat the number of economies contracting in a given year as if each economy contracted on its own schedule. Using World Bank data for twenty OECD countries, 1970 to 2023, this brief compares the real record with that independent benchmark. Contractions are far more clustered than independence allows: in 2020 nineteen of twenty contracted. The clustering does not change average growth, which is the same in both worlds. It shows up in three ways: the aggregate economy is $9\times$ more volatile than independence predicts, contractions are $2.3\times$ deeper when clustered ($p < 0.0001$), and the worst years carry about five percentage points of excess loss, which on the combined output of these twenty countries is roughly two to three trillion dollars per episode. The brief measures the impact, identifies what policy can act on, and says what it does not establish.

1 What was measured

Data. Annual real GDP growth (World Bank indicator NY.GDP.MKTP.KD.ZG) for twenty countries, 1970 to 2023: Australia, Austria, Belgium, Canada, Denmark, Finland, France, Germany, Italy, Japan, Korea, Mexico, the Netherlands, Norway, Spain, Sweden, Switzerland, Turkey, the United Kingdom and the United States. A country *contracts* in a year when its growth is below zero.

Independent benchmark. Each country contracts in a given year with its own observed frequency, independently of the others. The chance that exactly k of the twenty contract is then computed exactly, and compared with how many years actually had k contractions. This benchmark is the same, on average, as shuffling each country's growth series in time. The cached data and the script (`test_clustered_contractions.py`) are available from the author.

2 Contractions cluster

Figure 1 compares the number of years with each count of contracting countries against the independent benchmark. The real record has more years with none or few contractions than independence expects

(74% of years with two or fewer, against 55%) and far more with many: the variance of the count is 15.6, against 2.1 under independence.

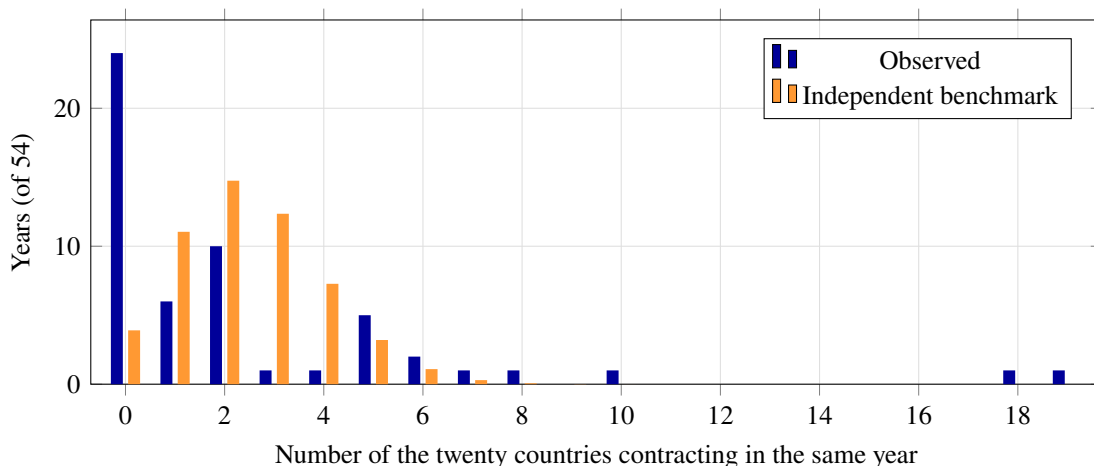


Figure 1: Years by number of contracting countries, 1970–2023. Independence expects almost every year to have between zero and five; the record has years with 7, 8, 10, 18 and 19.

Countries contracting	Years observed	Years expected	Independent benchmark: once in
at least 5	12	4.67	12 years
at least 7	5	0.374	144 years
at least 10	3	0.001882	28,695 years
at least 11	2	0.0002184	247,266 years
at least 19	1	1.4×10^{-15}	4×10^{16} years

The years are: at least 5, 1975, 1980–82, 1991, 1993, 2008–09, 2012–13, 2020 and 2023; at least 7, 1975, 1991, 1993, 2009 and 2020; at least 10, 1993, 2009 and 2020; at least 11, 2009 and 2020; 19, 2020 only. Two cautions. The upper rows rest on very few years: “at least 11” is two events, 2009 and 2020, so the exact figure for how rare they are under independence matters less than the fact that independence rules them out. And a count of contracting countries is not the same as a global recession: in 1975, 1991 and 1993 seven or more countries contracted, but the average growth of the twenty was still positive (+1.3% to +1.5%). Only 2009 (−3.6%) and 2020 (−4.3%) were broad, deep contractions.

3 The cost sits in the tail

Because each country keeps its own growth rates in the independent benchmark, average growth is identical in the two worlds: 2.69% a year (equal-weighted mean of the twenty countries). What differs is how the losses are distributed over years.

	Real record	Independent benchmark
Average growth	2.69%	2.69%
Worst year (mean of the twenty)	-4.25% (2020)	+1.23% (90% range +0.65% to +1.66%)
Mean of the five worst years	-0.98%	+1.60%
Years with seven or more contracting	5	0.38

In 5,000 draws of the independent benchmark, no draw had a worst year as bad as 2020's, and the draws' worst years ranged only from about +0.6% to +1.7%. A worst year that deep is what clustering produces.

4 How large is the economic impact

Three measures, each from the data.

The aggregate economy is 9× more volatile. The variance of the twenty-country average growth rate is 3.21 (standard deviation 1.79%), against 0.35 (standard deviation 0.59%) under independence. Clustering makes the world economy swing about three times further each year than it would if countries went up and down on their own schedules.

Contractions are 2.3× deeper when clustered. In the 62 country-year contraction episodes that fell in a clustered year (seven or more of the twenty contracting), the mean contraction was -3.09%. In the 70 episodes in non-clustered years it was -1.33%. A permutation test gives $p < 0.0001$, and within-country the pattern holds in 17 of 19 countries. The finding survives dropping 2020 (the ratio falls to 1.8×). A causal reading—that clustering itself deepens each country's downturn, through trade, financial contagion and confidence—is plausible but not proved by these data: global shocks may independently cause both clustering and depth.

Excess tail loss: about \$2 trillion per episode. In 2009, the twenty-country average fell -4.8 percentage points below where independence puts the worst year. In 2020 the gap was -5.5 points. On the combined GDP of these twenty countries, roughly \$40–50 trillion, each episode represents about \$2–3 trillion of output that fell because most economies contracted together, not one by one. Over the fifty-four-year record, the two episodes together amount to roughly \$4.7 trillion of excess tail loss—not an annual drag, but two catastrophic hits that an independence-based model would not anticipate.

What this does not measure. The long-run cost of clustering—lost investment, persistent unemployment, institutional damage—requires tracing output over the years after each episode. In these data, clustered contractions bounce back faster in the first two years (partly the 2020 COVID rebound) but show larger cumulative shortfalls at five years (1.6× the loss of non-clustered contractions). With only five clustered episodes, this is suggestive, not conclusive.

5 At the other end, independence overestimates

For one or two contracting countries, the real record has fewer years than independence predicts (the ratio of observed to expected is about 0.5 to 0.7). At three or four the record has far fewer than expected (one year each, against twelve and seven expected). Contractions bunch into a few years, and 24 of the 54

years had none at all. The pattern is the sign flip described in the book: independence is too pessimistic about common, modest events and far too optimistic about rare, severe ones.

6 Why this brief reports counts and not a single coupling constant

The book's law, $R \approx d^n$, would suggest fitting one number d to how fast the ratio of observed to expected years grows with the number of contracting countries. For this record the ratio is not a straight line on a log scale: it is about 0.6 for one to three countries, 1.1 at four, 2.6 at five, 13 at seven, 1,600 at ten, and keeps climbing because the independent benchmark falls away faster than the data do (two years at every count from 11 to 18). A fitted d therefore depends on how many years must support each count: about 1.9 if each count must be seen in at least five years, 2.7 at three years, and 6.8 or more if single events are allowed (90% bootstrap ranges are wide, for example 1.8 to 7.0 at three years). A single d is not a stable summary of this record. The data look more like two regimes, ordinary years and a few years in which nearly every economy contracts together.

7 Before and after 1990

Clustering appears in both periods. In 1970–1989, one year had seven or more of the twenty contracting (independence expects 0.02); in 1990–2023, four did (independence expects 0.49), and all three of the years with ten or more contracting fall after 1990. With five events in all, this suggests more frequent severe clustering since 1990 but cannot establish it.

8 Policy actions

What is at stake. Reducing clustering does not raise average growth, which is the same either way. It compresses the distribution toward the mean. In a world where each economy's contractions fell independently of the others, the worst year over a half-century would average about +1.2% growth instead of −4.3%—a swing of roughly five and a half percentage points. The five worst years would average +1.6% instead of −1.0%. No country loses any growth—each still has all its boom years, just not at the same time as every other country. The twenty-country average in the best year drops from +6.2% to about +4.0%, because some booms are also synchronized, but that is a change in timing across countries, not a loss of output. The median year barely moves (+3.1% to +2.7%). The gain is concentrated in the tail: the worst years improve by five and a half points while the best-year average falls by about two. The year-to-year volatility of the global economy falls by about two thirds (standard deviation from 1.79% to 0.59%). No policy can fully decouple twenty economies, but even a partial reduction in clustering would narrow the gap.

The measurements in this brief point to five actions, ordered from what any institution can do today to what requires coordination.

1. Measure the clustering. Run Figure 1 for any system where independence is assumed: count how many components fail simultaneously, compare with the exact independent benchmark, and publish the ratio. If observed failures at the tail exceed the benchmark by an order of magnitude or more, the

risk model is wrong where it matters most. This is a single script on existing data. *Evidence: the observed/expected ratio reaches 1,600× at ten or more countries (Section 2).*

2. Stress-test for the clustered scenario. Post-2008 regulatory stress tests do include multi-economy downturns, but the question is whether they test *enough* simultaneous contraction. The record shows that in the worst years, 18 to 19 of the 20 economies contract simultaneously. If no stress scenario goes that far, the test misses the actual tail. Regulators should require at least one scenario in which most major economies contract together, calibrated to the 2009 or 2020 episode. *Evidence: 19 of 20 contracted in 2020; under independence this happens once in 4×10^{16} years (Section 2).*

3. Size fiscal buffers to the clustered tail. A counter-cyclical buffer sized to the average contraction (−1.3%) is exhausted by a clustered one (−3.1%, and −4.3% at the extreme). Sovereign wealth funds, automatic stabilisers and emergency lending facilities should be sized to the worst year in the record, not to the average downturn. *Evidence: clustered contractions average −3.1% against −1.3% for non-clustered; the worst episode hit −4.3% (Sections 3–4).*

4. Make coordination mechanisms standing, not improvised. The 2009 G20 fiscal stimulus and the 2020 central-bank swap lines worked but were assembled under crisis pressure. Standing arrangements—automatic swap lines triggered when a threshold count of economies enters contraction, pre-agreed fiscal coordination frameworks—would deploy faster and with less political friction. The data show that when seven or more economies contract together, each contraction is 2.3× deeper on average. The aggregate economy is already 9× more volatile than independence predicts across the full record; pre-committed coordination would reduce the tail that drives most of that volatility. *Evidence: 2.3× deeper contractions when clustered ($p < 0.0001$); 9× overall volatility (Section 4).*

5. Diversify supply chains as macroprudential policy. This brief cannot measure how much trade concentration drives clustering (Section 9), but the logic is direct: if the same shock hits all of a country’s trade partners simultaneously, exports, imports and capital flows all fall at once. Supply-chain diversification is not only a growth strategy—it is a hedge against synchronized downturns. *Evidence: not established for this data set; the trade channel is untested (Section 9).*

9 What this brief does not establish

- That trade structure drives the clustering. An earlier draft compared a trade-concentrated group of five economies with a group of fifteen and reported a threefold difference in coupling. The fitted coupling depends on group size, so that comparison was an artifact; a fair test needs size-matched groups and verified trade data, which have not been done.
- Any effect on income distribution within countries.
- That any particular institution’s model assumes independence. The point is about what the independent benchmark implies, not about what any named model does.

10 Reproducing this

All computations use `test_clustered_contractions.py` (benchmark, bootstrap, worst-year draws) and `build_clustered_brief.py` (volatility, depth, within-country, permutation test, best-year com-

parison; generates this document). Scripts and data: available from the author. Full method: roguewave-book.com.