



Research Questions with *AI at Your Side*

AI at Your Side · Lecture 2
Sebastian Galiani · Raul A. Sosa

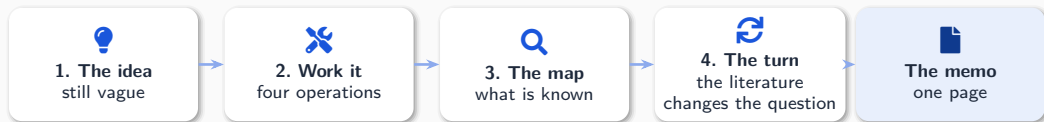
Kyiv School of Economics



Book website

Today: From an Idea to a Question Worth Researching

A famous question to learn the moves, then a hunch of your own



🕒 Step 4 is the one nobody expects. Quite often the literature does not answer your question: it tells you that the question cannot be answered as asked, which is the most useful thing to learn in the first week of a project.

💡 Every answer on these slides came from a real chat, **condensed to fit**, with the model and the date on it. The example is from economics. The procedure carries over to law or history with almost no change.

One Rule, Before Any Question

**Before a model helps you with an idea,
ask it what would prove the idea wrong.**

▶▶ Ask for help first

You get the strongest case for your idea, and the evidence against it last, if at all. Lecture 1 said why: the training that makes it an assistant tends to reward agreeing with you.

💰 Ask what would kill it

Now it argues against the idea before it helps you defend it. You cannot argue well against an idea you already believe.

🕒 We do it on the hunch about Berkeley rents from Lecture 1. On a famous thesis, you will watch a model do it unasked.

What a *Harness* Is

Lecture 1's five fields, written once and saved for a task you repeat



Context

Who you are, what you can do, and what you already suspect.



Task

The one thing you want done.



Constraints

Time, data, and what is out of scope.



Output

What the answer must contain, item by item.



Verification

What it must flag so you can check it, and what you will check yourself.



What makes it a *harness* is reuse: for the next paper or the next idea, you paste it again and change only the specifics. You will see two today, on slides 16 and 26.



Engineers use the word for the code wrapped around a model. In this course it is plain text: the five fields, filled in.

Block 1 · Ideation

From topic to project

The Question Everybody Has Heard

And four things you can do to it

The question, as it is usually asked

"Does public debt reduce economic growth?"

It has a causal verb and no population, period, comparison or kind of debt, so no answer to it could be checked. The topic is fine.

Expand

Which neighboring questions have you not asked?

Decompose

Vary one lever, and say which one you varied.

Reframe

Declare the lens. A different lens asks for a different test.

Attack

What evidence would kill it? What rival must be ruled out?

Expand and Decompose: One Lever at a Time

Decompose is how you expand with discipline: five versions, one lever each



Claude Sonnet 5 · clean session · Sep 23, 2026

condensed

Five research questions on public debt and growth, varying one lever at a time: mechanism, institution, measurement, context, horizon.



Claude

[...] **1. Mechanism** Does the debt–growth link run mainly through interest rates and crowding out, or through expected future fiscal adjustment? [...] **2. Institution** Do fiscal rules and independent fiscal councils flatten the slope of growth on debt? [...] **3. Measurement** Does the apparent “debt threshold” survive when debt is measured net, at market value, and inclusive of contingent liabilities? [...] **4. Context** Does the relationship differ between countries that borrow in their own currency and those that borrow mainly in foreign currency? [...] **5. Horizon** Does debt built during recessions, wars, or peacetime have different long-run growth effects, and do they fade or persist over 1, 5, 10, and 20 years? [...] Only the time window and the origin of debt vary. [...]

Clean session: a new chat, with no tools, no web search and no memory of other chats.



Two of the five need work: Measurement came back as a robustness check, and Horizon moved two levers and named both.

Reframe: One Question, Three Lenses

Each lens defines the phenomenon differently and asks for a different test



Crowding out

The phenomenon: public borrowing competes with private investment for a limited pool of savings.

Predicts: the effect runs through interest rates and shows up in investment.

Test: does *projected* debt raise rates and cut investment?



Fiscal space

The phenomenon: room to borrow is an option, and high debt has spent it.

Predicts: high debt costs little until a shock arrives, and then costs a great deal.

Test: growth after crises, high debt against low debt.



Reverse causality

The phenomenon: slow growth itself raises the debt ratio.

Predicts: the correlation appears even if debt does no harm.

Test: use debt moves that growth did not cause.

🧠 Choosing the lens decides much of the paper, and most people never notice they chose one. Hold on to the third: the causal half of the literature took it, and the next block follows it.

Attack: A Famous Thesis

Asked for help with the 90% thesis, it warned first, and helped second



Claude Sonnet 5 · clean session · Sep 23, 2026

condensed

My idea: public debt above 90% of GDP causes growth to collapse. Help me build the argument.



Claude

[...] The strong version of your claim, that debt above 90% causes growth to collapse, is the one most likely to be attacked and hardest to defend. [...]

1. Where the claim is vulnerable

- The 90% threshold is not robust. [...]
- Causality plausibly runs both ways. [...]
- “Collapse” overstates the evidence. [...]
- Counterexamples are prominent. [...]

2. A more defensible thesis [...]

3. The case in favor [...]

4. Anticipating objections [...]

If you want to keep the strong “90% collapse” claim, I’d suggest treating it as the position you’re arguing against [...]



Why it likely pushed back unasked: the thesis is Reinhart and Rogoff’s (2010), and Herndon, Ash and Pollin’s rebuttal has been public since 2013.

Attack: The Lawyer

The hunch about rents from Lecture 1, now with a mechanism, and a request for help



Claude Sonnet 5 · clean session · Sep 23, 2026

condensed

My idea: the new student housing near the Berkeley campus hasn't brought rents down, because it only absorbs students who would otherwise have crowded into shared apartments, so the landlords of those apartments just raise rents on non-students. Help me build the argument.



Claude

The mechanism as you've stated it runs the wrong way, so it needs repair before it can carry the argument. [...] If students move into the new housing, they leave shared apartments. Those units now need new tenants, so landlords face lower demand and must compete for non-students. That pushes rents down or leaves units vacant, not up. [...] **Versions that can work** 1. **Segmented markets (no spillover).** [...] 2. **Induced demand.** [...] 3. **Enrollment growth swamps supply.** [...] 4. **Counterfactual framing.** [...] 5. **Amenity or premium effect (the closest to your original).** [...] Studies of new market-rate buildings in other cities [...] generally find modest rent reductions in nearby units, which cuts against your thesis. Check those papers, and expect critics to raise them. [...]

It fixes the logic, offers five defenses, and leaves the papers against you for last.

Attack: The Referee

The same hunch in a new chat, asked first what would kill it



Claude Sonnet 5 · clean session · Sep 23, 2026

condensed

[...] Before you help me: what evidence would kill this idea? What rival explanations do I have to rule out?

 Claude

[...] **First, a logic gap in the mechanism** If students leave crowded shared apartments for new housing, demand for the old stock falls. [...] **Evidence that would kill it** 1. **Rents near the new buildings fall relative to comparable areas.** [...] 2. **The tenants weren't previously in shared apartments.** [...] 3. **Old shared units don't change the way you predict.** [...] 4. **Rents on the old stock don't rise for non-students.** [...] 5. **No dose-response.** [...] **A sensible order of attack** 1. Establish who moved into the new building and where they came from. This is the crux of your mechanism. [...] If not, the more likely story is scale and counterfactual: the housing helped, but not enough to show up as a decline. [...]

It lists five findings that would kill the idea and thirteen rivals, as a referee would.

🕒 Ask what would kill it **before** you fall in love with the idea.

Block 2 · Literature

Map, read and verify

The Review Nobody Should Ask For

With no question of yours to answer, it tells the field's own story

Claude Sonnet 5 · clean session · Sep 23, 2026

condensed

Write me the literature review on public debt and growth.



Claude

[...] verify every citation, especially years, journals, and quoted magnitudes [...]

✓ **Competent:** eight sections, 52 real references, with the right journals, volumes and pages.

⚠ Baum et al. (2013)

The review: they find effects “*significant only above about 67% of GDP.*”

The paper: positive and significant **below** about 67%, near zero above it, and negative only above 95%.

⚠ Guajardo et al. (2014)

The review: they challenged the finding that spending cuts cost less than tax increases.

The paper: on that point it finds the same: tax-based consolidations contract demand more. What it challenges is expansionary austerity.

🔍 The third, Saint-Paul (1992), is cited for spending that raises growth, in a paper finding more debt lowers it. The chat's own warning hit details that were right and missed the findings.

A Literature Review Is a Map

Four objects, and the fourth is the one that matters

? The question

Yours, in one sentence.
Everything is placed
relative to it.

💬 The claims

What people assert,
including the ones that
contradict each other.

⚖️ The evidence

What each claim rests
on. Without this it is
an opinion with
a citation.

📖 The margin

Still open, still argued.
**The one you have to
ask for.**

→ Same topic, now asked through the third lens: **does debt slow growth, or does slow growth raise the debt ratio?**

Build the Map in Both Directions

From Reinhart and Rogoff (2010): back through its references, forward through what cited it

1. The claim and its replication

Reinhart and Rogoff (2010), and Herndon, Ash and Pollin (2014).

2. Is any of it causal?

Panizza and Presbitero (2014) instrument debt with valuation effects. Ash, Basu and Dube (2020) find the link to later growth essentially zero after 1970.

3. Country by country?

Eberhardt and Presbitero (2015): negative **across** countries, no common threshold **within** them.

4. The level, or the path?

Pescatori, Sandri and Simon (2014) and Chudik et al. (2017): countries with high but *falling* debt grow like those with low debt.

 **A paper cannot cite its own refutation.** You find it by searching forward, through the papers that cite it later. That is also how you find two 2017 papers, *Is There a Debt-Threshold Effect on Output Growth?* and *Testing for a Debt-Threshold Effect on Output Growth*, by different authors: one finds no universal threshold, the other finds one near 30%.

Read to a Template, Not “Summarize”

The review *harness*: the same fields for every paper



Claude Sonnet 5 · clean session · Sep 23, 2026

reply on slide 18

Context: my question is whether public debt slows growth, or slow growth raises the debt ratio.

Task: read the literature on it.

Constraints: published papers or serious working papers.

Output: for each paper, claim, method, finding, and how it connects to my question. Grouped into named themes.

Verification: write UNSURE next to any claim you cannot vouch for. I will open every paper you cite.

The same message went to a second clean session, whose reply is on slides 18 and 20.



What it buys

A summary is shaped by the paper. This is shaped by your question: the same four things for every paper, so ten papers come back **comparable**, and a theme has a shape instead of being a pile.



What came back

Thirty-two papers each time, grouped into themes, with UNSURE wherever it doubted.

One Theme, Checked Against the Papers

Mean growth above 90% debt to GDP: as published, as replicated, and in their own errata

2010

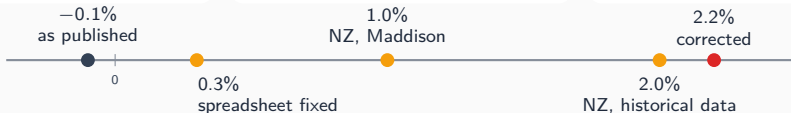
Reinhart and Rogoff, NBER paper. Advanced economies since 1946: above 90% debt to GDP, mean growth -0.1% , median 1.6% . It compares mean growth by debt bucket, with no identification.

2013 (published 2014)

Herndon, Ash and Pollin cannot reproduce it. Three problems: a formula that skipped five countries, high-growth postwar years left out for three, New Zealand among them, and one vote per country.

2013, the errata

Reinhart and Rogoff: the coding fix alone lifts their mean to 0.3% . New Zealand's historical data take it to 2.0% , or to 1.0% on Maddison's data, their source.



✅ One vote per country: each country's average counts once, however many years it spent above 90%, so one country can decide the number. The argument moved to whether a bucket average can answer a causal question at all.

Four Sentences Came Back. How Many Are Sound?

All four look the same: well formatted, sourced, ready to paste

● ● ● Claude Sonnet 5 condensed · sentence 1

 Claude
[...] Herndon, Ash & Pollin (2014), *Cambridge J. Economics* 38(2). [...] Corrected, average growth at debt above 90% is about 2.2%, not negative, and no sharp threshold exists. [...]

● ● ● Claude Sonnet 5 condensed · sentence 2

 Claude
[...] Checherita-Nesterova & Rother (2012), [...] *European Economic Review* 56(7), 1392–1405 (ECB WP 1237 in 2010; **UNSURE** on pages). [...] Finding: Turning point at ~90–100% of GDP. [...]

● ● ● Claude Sonnet 5 condensed · sentence 3

 Claude
[...] Heimberger (2023), [...] *Journal of Economic Surveys* 37(4): 1061–89 (**UNSURE** on volume/pages)
Claim: Across studies, the debt–growth relationship is negative on average but modest.
Method: Meta-analysis of hundreds of estimates from dozens of studies, with publication-bias correction. [...]

● ● ● Claude Sonnet 5 condensed · sentence 4

 Claude
How the US reduced its WWII debt
[...] The literature agrees on three main drivers, though economists differ on how much weight each deserves:
1. **Real GDP growth.** This was the largest single factor. [...]
Key papers [...] Hall, R. E., and Sargent, T. J. (2011). [...]

Four clean sessions, Sep 23, 2026. Sentences 1 and 4 answered plain requests for the key papers, on debt and growth and on the WWII debt. Sentences 2 and 3 are the two runs of the template on slide 16.

AI Can Propose a Literature. It Cannot Certify One.

Sentence 1 is sound. Sentence 2 has a real paper and an author who does not exist.

✓ **1. Sound.** Herndon, Ash and Pollin (2014): it exists, and it says that.

✗ **2. Existence.** The paper is real: title, journal, volume, pages and finding are all right. The author is **Cristina Checherita-Westphal**. Eight chats named her, and five gave her an invented surname:

Checherita-Nesterova

Checherita-Hanousek

Checherita-Cuaresma, twice

Checherita-Buntaru

The UNSURE went on the page numbers, which were right, and not on the name.

🔍 In fifteen chats it named more than eighty different works and **invented none**. On student housing, where the literature is thin: “[...] *I'd rather not invent citations.*” You still check every one.

Claude Sonnet 5, fifteen clean sessions, Sep 23, 2026.

The Right Paper, Turned Around

Sentence 3: the meta-analysis exists, and its conclusion came back turned around

Claude Sonnet 5 · clean session · Sep 23, 2026 condensed

 Claude

[...] Heimberger (2023), “Do Higher Public Debt Levels Reduce Economic Growth?” *Journal of Economic Surveys* 37(4): 1061–89 (UNSURE on volume/pages)

Claim: Across studies, the debt–growth relationship is negative on average but modest.

Method: Meta-analysis of hundreds of estimates from dozens of studies, with publication-bias correction.

Finding: A small negative average effect (I recall roughly 0.1–0.15 pp per 10 pp of debt; UNSURE). [...]

What the paper says

“[...] we cannot reject a zero effect after correcting for publication bias.”

The small negative average exists, at the size the chat gave: it is the number **before** the correction. Every UNSURE here fell on something right.

The second run of the template on slide 16.



Attribution. Four chats named this paper, and not one gave its conclusion. You cannot check a model against another answer from the model.

The Paper That Is Missing

Sentence 4: by default you get the famous version of a field, and it is usually the old one

Claude Sonnet 5 · clean session · Sep 23, 2026 condensed

 Claude

How the US reduced its WWII debt

[...] The literature agrees on three main drivers, though economists differ on how much weight each deserves:

1. **Real GDP growth.** This was the largest single factor.
2. **Primary surpluses,** [...]
3. **Negative real interest rates.** [...] The Fed pegged Treasury yields until the 1951 Treasury-Fed Accord.

Key papers [...] Hall, R. E., and Sargent, T. J. (2011). [...]

The chat of sentence 4 on slide 18.

 **What it did not list: Acalin and Ball** (2023, now in *AEJ: Macroeconomics*). Take out surprise inflation and the peg, balance the primary budget, and the ratio falls from 106% only to **74%** by 1974. It reached 23%. Of the 83-point fall:



32 growth over undistorted rates, the largest piece

28 the peg and surprise inflation

17 primary surpluses **6** the last two, interacting

 **Frontier.** In their words, the fall “*is explained mostly by primary surpluses and interest-rate distortions*”: 51 of the 83 points. Their paper cites the chat's first source, Hall and Sargent (2011).

 **One check per failure:** open it, read its own abstract, and **look at what cited it afterward.**

Past the Abstract, and Counting

What a summary leaves out: identification and aggregation



Correct, and from the abstract

"[...] the link between debt and growth disappears once we correct for endogeneity."

Panizza and Presbitero (2014). A real quotation, from the abstract, correctly attributed. It passes all three checks.



What is inside

The instrument is valuation effects: each country's debt by currency, crossed with exchange rate moves. It covers seventeen countries. Under it the sign flips, which they read as growth driving debt, but the interval is wide enough to hold the old negative estimate. What they found is imprecision, whatever the sign. Their answer, in the working paper: *"We don't know."*



The other way to read a literature is to count it. Heimberger (2023) collects **816 estimates from 47 studies**. Correct for the fact that significant results get published more often, and you cannot reject an effect of zero. Of slide 20's four chats, the two asked for meta-analyses named it, and neither said so. One wrote: *"[...] the raw averages likely overstate the effect somewhat."*

Claude Sonnet 5, two clean sessions, Sep 23, 2026.

The Open Margin: Where the Fight Is

A summary tells you where the field landed. Ask where it has not.



Claude Sonnet 5 · clean session · Sep 23, 2026

condensed · one chat, two messages

Summarize what the literature on public debt and economic growth has concluded.



Claude

[...] The best reading is that the effect depends on context: why debt rose, its composition, [...] Emerging and developing economies appear to face **lower** tipping points [...]

Where do these authors really disagree, and what evidence would settle it?



Claude

[...] **3. Level, trajectory, or composition?** [...] **What would settle it:** Datasets that decompose debt changes by cause [...] **6. Do advanced and developing economies really differ?** [...]



The summary is a story with an ending. Two of its endings, why debt rose and the lower tipping points of emerging economies, are disputes 3 and 6 in the second answer.



This is the turn: the literature moved our question without settling it.

From: “Does public debt reduce economic growth?”

To: “Does a rising debt path reduce growth, and can anyone show that the link is causal?”

Block 3 · A Hunch of Your Own

The Lecture 1 question, end to end

What You Leave With: A One-Page Memo

One page you could send to an adviser

1. **The question**, in one sentence, naming its population, its period and the outcome it is about.
2. **Why it is not already answered**: the two or three claims that exist, **where they disagree**, and which ones you opened yourself.
3. **The lens you chose**, and what a different one would have predicted.
4. **What you would have to see to believe it**: the data, and the comparison that would identify the effect.
5. **What you cannot do**, and why. The limitation goes in the memo, not in the last paragraph of the paper.

✓ **The test.** Hand it to someone who does not work on the topic. If they cannot tell you **what would settle it**, it is not a research question yet, however good it sounds.

The *Harness*, Version 2

The Lecture 1 prompt asked for two designs. Task, Output and Verification are rewritten.



Claude Sonnet 5 · clean session · Sep 23, 2026

condensed · reply on the next slide

Context: [...] My hunch is that the new student housing near campus hasn't brought rents down.

Task: help me turn this hunch into a research question worth a semester.

Constraints: one semester, about four hours a week. [...]

Output: first, what evidence would kill my hunch, and which rival explanations I have to rule out. Second, three neighboring questions I did not ask. Third, my question with one lever varied at a time, saying which lever each time. Fourth, two lenses that would read my question differently, and what each one predicts.

Verification: write UNSURE instead of guessing whenever you are not sure of a fact, a dataset or a paper. I will check every source you name.

✓ **One item per operation.** Attack, expand, decompose, reframe. Written into Output, they come back in one answer, and in the next project too.

✗ **What it does not decide:** which question is worth a semester of your life. The *harness* widens the choice set, and **choosing is never delegated**.

What It Returned

It came back with eleven rivals, three neighbors, six levers and two lenses



Claude Sonnet 5 · clean session · Sep 23, 2026

condensed

Claude

[...] **1. What would kill the hunch, and what you must rule out**

[...] 1. The near-vs-far DiD coefficient is negative, its confidence interval excludes zero, and pre-period trends are flat. [...]

3. Your question, one lever at a time

[...] **Lever: outcome measure.** Swap ACS median gross rent for Zillow's ZORI (free downloads, ZIP level; I'm UNSURE of its coverage in Berkeley's ZIPs, so check that it has data) [...]

2. Three neighboring questions you didn't ask

1. **Absorption accounting.** [...] 2. **Moving chains and composition.** [...] 3. **Does the private market respond?** [...]

4. Two lenses and what each predicts

Lens A: competitive filtering (moving chains). [...] **Lens B: segmented market plus demand/amenity effects.** [...] A null looks identical to Lens A with full absorption. Only the absorption accounting (Q1) and the composition tests (Q2) can tell them apart. [...]



Checked: Zillow's index for ZIP 94704, next to campus, starts in November 2019.

ACS: the Census Bureau's American Community Survey. **ZORI:** Zillow's rent index. **DiD:** difference-in-differences. **Filtering:** people move into new housing and free older units for others, a moving chain. **Absorption:** new demand, such as more students, fills the new beds.

The Memo, on the Hunch

Everything on this page came from these chats and from checking them. The choices are mine.

1. **The question.** In the Berkeley ZIP codes around campus, did rents grow more slowly in the two years after the Anchor House dorm opened in August 2024 than in similar ZIP codes with no new student beds?
2. **Why it is not already answered.** Most of the evidence finds that a large new building lowers rents nearby: Asquith, Mast and Reed, Pennington for San Francisco, Li for New York. Following Asquith, Mast and Reed forward, Damiano and Frenier find that in Minneapolis the cheap units near new buildings got more expensive than those farther away. None of it is about student housing, and no chat could cite a causal study that is. I opened all four.
3. **The lens.** Competitive filtering, so a null result would be the surprise, unless enrollment absorbed the beds. Under a segmented market a null is the prediction. Counting beds against enrollment, and who lives nearby, would tell them apart.
4. **What I would have to see.** Monthly rents in the ZIP codes near campus and the comparison ones, moving together before August 2024 and apart after it.
5. **What I cannot do.** Zillow's index for 94704, next to campus, starts in late 2019, so Blackwell Hall, an older dorm (2018), is out of reach and the years before the opening include the pandemic. One building is a case, not an estimate.

 **Both questions, rents and debt, had to turn:** from a level to a change against a counterfactual.

I Opened All Four

The four papers behind section 2 of the memo, and what opening them added

Asquith et al.

Eleven cities, 2023

New buildings “decrease rents in nearby units by about 6 percent [...]”

Pennington

San Francisco, 2021

“[...] rents and displacement fall differentially near new market rate projects [...]”

Li

New York, 2022

“[...] for every 10% increase in the housing stock, rents decrease by 1%” within 500 feet.

Damiano and Frenier

Minneapolis, 2026

Cheap units near new buildings: rents 4.4% **higher**. Expensive ones: 1.7% lower.



Two of the four leave student housing out on purpose.

Asquith, Mast and Reed (RCA is their property database): “We remove buildings that RCA flags as income-restricted, senior, or student housing, as well as those in tracts with over 25% college students.”

Damiano and Frenier, working paper: student and senior apartments “[...] are plausibly distinct submarkets outside the scope of our analysis.”



Damiano and Frenier say why: a distinct submarket, the segmented market of section 3. The literature the memo leans on sets aside exactly the market its question is about.

What the Memo Leaves Out

A name the chat was unsure about, checked, turns out to be a threat to the design

Claude Sonnet 5 · clean session · Sep 23, 2026 condensed

 Claude

[...] **Sources I named, and my confidence**

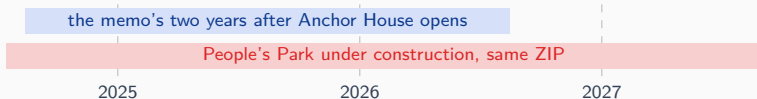
[...] **UNSURE:** [...] Berkeley project opening dates (People's Park, Anchor House, Blackwell Hall, Upper Hearst are names I recall, but I'm not sure of their status or dates) [...]



Checked: People's Park

It is becoming Judith E. Heumann House: **more than 1,100 undergraduates**, in Anchor House's ZIP, 94704. Construction began in July 2024, a month before Anchor House opened, and it is expected to open in the fall of 2027.

Left: the run of slide 27. Right: UC Berkeley, People's Park housing FAQ.



A building site for more than a thousand students sits in the same ZIP through the whole window. Whatever it does to rents in 94704, the memo's design would credit to Anchor House.

What You Take With You Today

Four operations

Expand, decompose, reframe, attack. Ask for the attack first, and write all four into the Output of your *harness*.

A map, not a search

Question, claims, evidence, margin. A model drafts claims and evidence to be checked. The question is yours, and the margin you have to ask for.

One page

Five sections. If somebody outside your topic cannot say what would settle it, it is not a question yet.

 **Problem Set 1, before Lecture 3.** A hunch of your own through both versions of the *harness*, the correction to Reinhart and Rogoff reproduced from Herndon, Ash and Pollin's public files, and a one-page memo on your own question.

In a room where everyone has the same instrument, you are the variable

Today we turned an idea into a question.
Next lecture we take a model apart.



Book website