

# Playing telephone with generative models: “verification disability,” “compelled reliance,” and accessibility in data visualization



Frank Elavsky, PhD Candidate



Human-  
Computer  
Interaction  
Institute



hcii.cmu.edu, axle-lab.com, dig.cmu.edu

2025

[frank.computer](http://frank.computer)

# (this workshop paper could have been a blog post)



Frank Elavsky, PhD Candidate



[hcii.cmu.edu](http://hcii.cmu.edu), [axle-lab.com](http://axle-lab.com), [dig.cmu.edu](http://dig.cmu.edu)





Sep 30, 2025, 10:38 PM (24 days ago)



my thoughts? gut reaction? lmao not your best work. probably should have been a blog post.

you're right that there's not much at this specific intersection. but maybe address this head-on? like: "Yes, we know AI hallucinates. Yes, we know accessibility matters. But here's why the combination creates a novel problem space that existing work doesn't address." does that help? o think the work is solid and important - it just needs to position itself as "applying known problems to an underexamined context" rather than "discovering unknown problems." the contribution is the framing and the specificity, not the existence of ai bias itself.

**A good workshop paper  
stimulates discussion.**

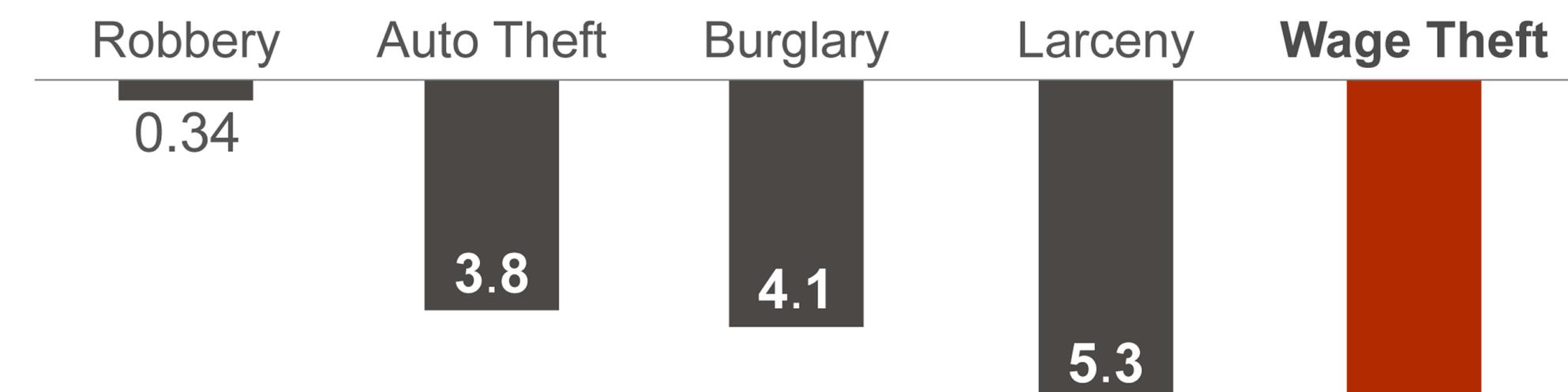
# A good workshop paper stimulates discussion.

I'll try to have this talk make up for where the  
workshop paper falls short!

**We decided to play a game of  
telephone...**

# Money Lost Each Year to Theft

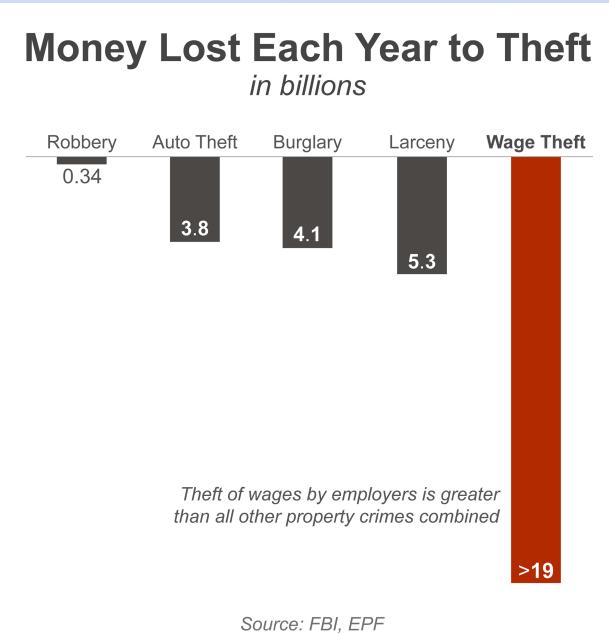
*in billions*



*Theft of wages by employers is greater than all other property crimes combined*

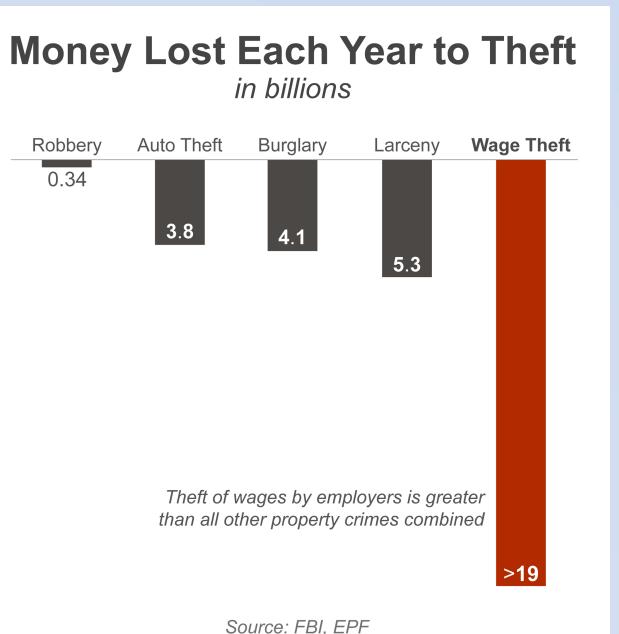
Source: FBI, EPF

# chatGPT (1st instance)



“Can you describe this?”

# chatGPT (1st instance)



“Can you describe this?”

Certainly. The image is a bar chart titled “Money Lost Each Year

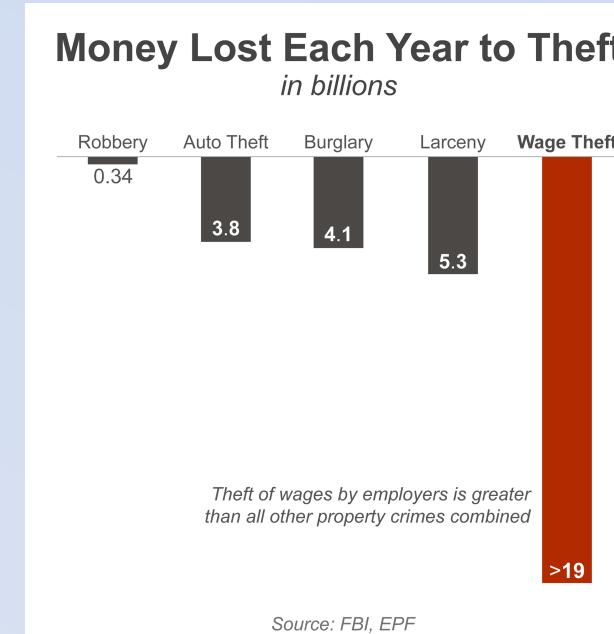
to Theft (in billions)”. It compares different types of theft based on  
the estimated amount of money lost annually in the United States.

The data source is attributed to FBI and EPF.

Breakdown of the Chart:

- Robbery: \$0.34 billion
- Auto theft: \$3.8 billion
- Burglary: \$4.1 billion
- Larceny: \$5.3 billion
- Wage Theft: > \$19 billion (highlighted in red)

# chatGPT (1st instance)



## “Can you describe this?”

Certainly. The image is a bar chart titled “Money Lost Each Year to Theft (in billions)”. It compares different types of theft based on the estimated amount of money lost annually in the United States.

The data source is attributed to FBI and EPF.

Breakdown of the Chart:

- Robbery: \$0.34 billion
- Auto theft: \$3.8 billion
- Burglary: \$4.1 billion
- Larceny: \$5.3 billion
- Wage Theft: > \$19 billion (highlighted in red)

Key Message:

At the bottom, the chart states:

“Theft of wages by employers is greater than all other property crimes combined.”

This emphasizes that wage theft—when employers illegally withhold wages or underpay workers—is a significantly larger issue financially than traditional forms of property crime.

The chart uses color and size to underscore this disparity, with “Wage Theft” shown as a much taller red bar compared to the smaller gray bars for other crimes.

# chatGPT (2nd instance)

“Can you make a chart based  
on this description?”

Certainly. The image is a bar chart titled “Money Lost Each Year to Theft (in billions)”. It compares different types of theft based on the estimated amount of money lost annually in the United States.

The data source is attributed to FBI and EPF.

Breakdown of the Chart:

- Robbery: \$0.34 billion
- Auto theft: \$3.8 billion
- Burglary: \$4.1 billion
- Larceny: \$5.3 billion
- Wage Theft: > \$19 billion (highlighted in red)

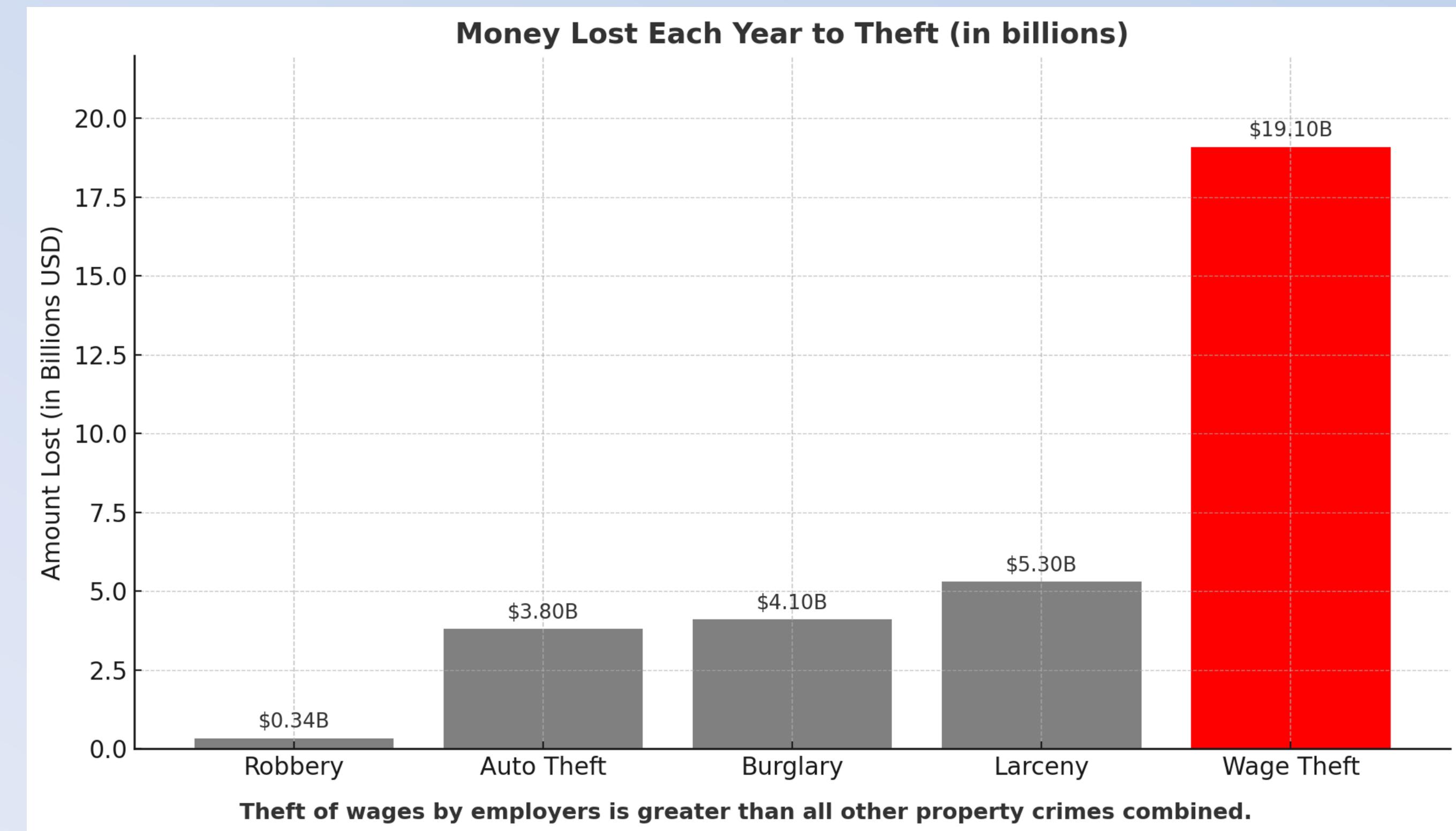
Key Message:

At the bottom, the chart states:

“Theft of wages by employers is greater than all other property crimes combined.”

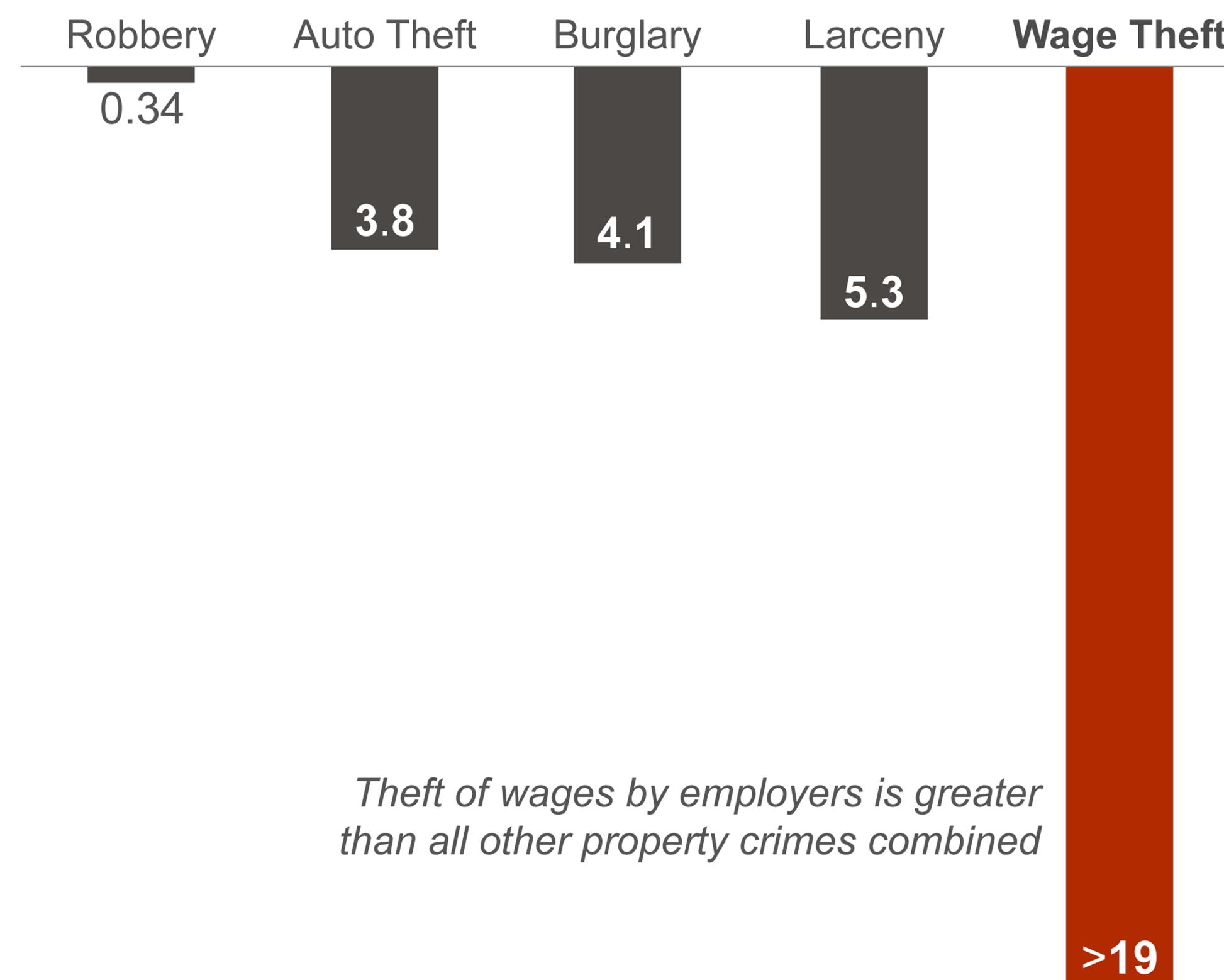
This emphasizes that wage theft-when employers illegally withhold wages or underpay workers-is a significantly larger issue financially than traditional forms of property crime.

The chart uses color and size to underscore this disparity, with “Wage Theft” shown as a much taller red bar compared to the smaller gray bars for other crimes.

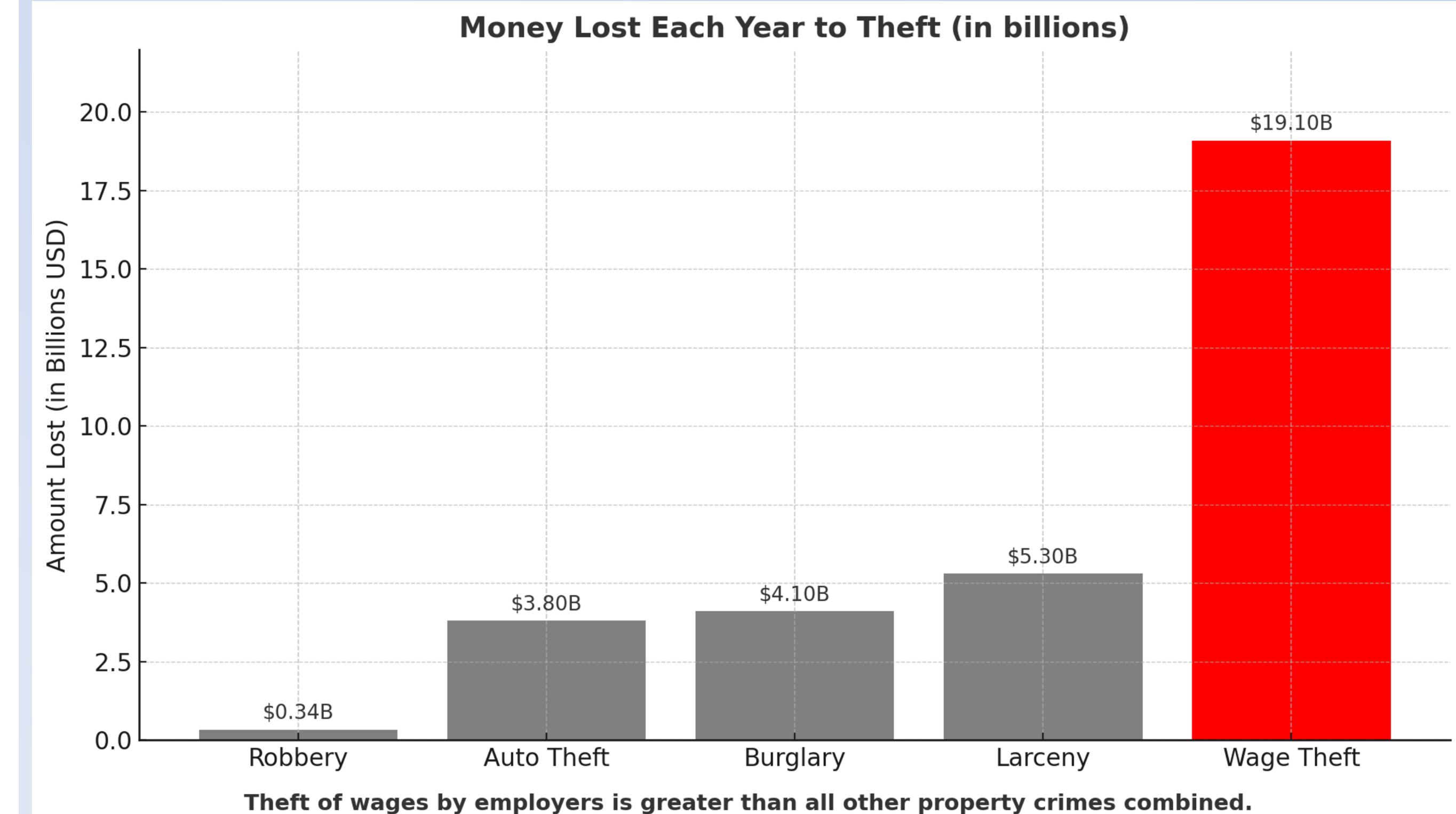


# Quick comparison, so far

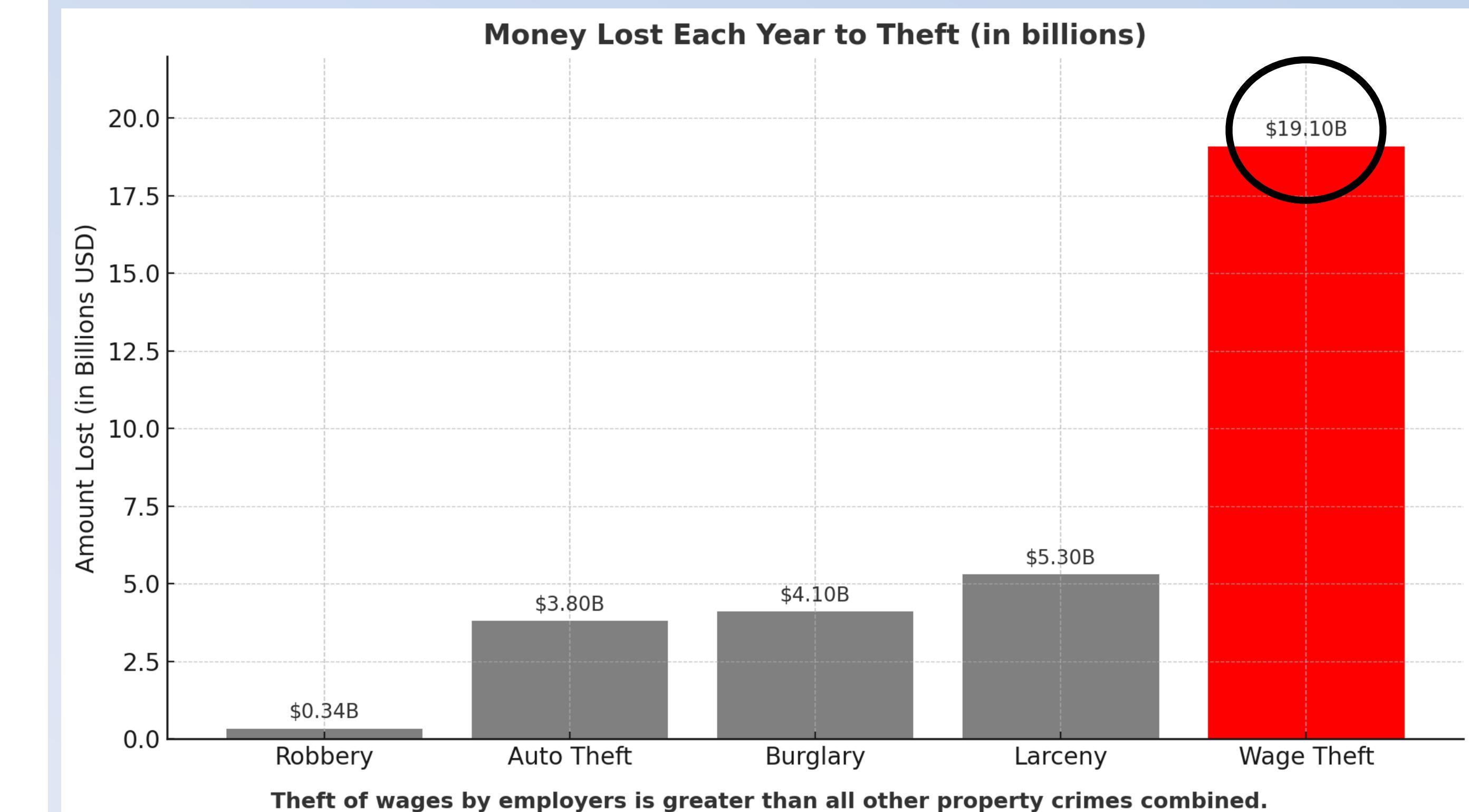
## Money Lost Each Year to Theft *in billions*



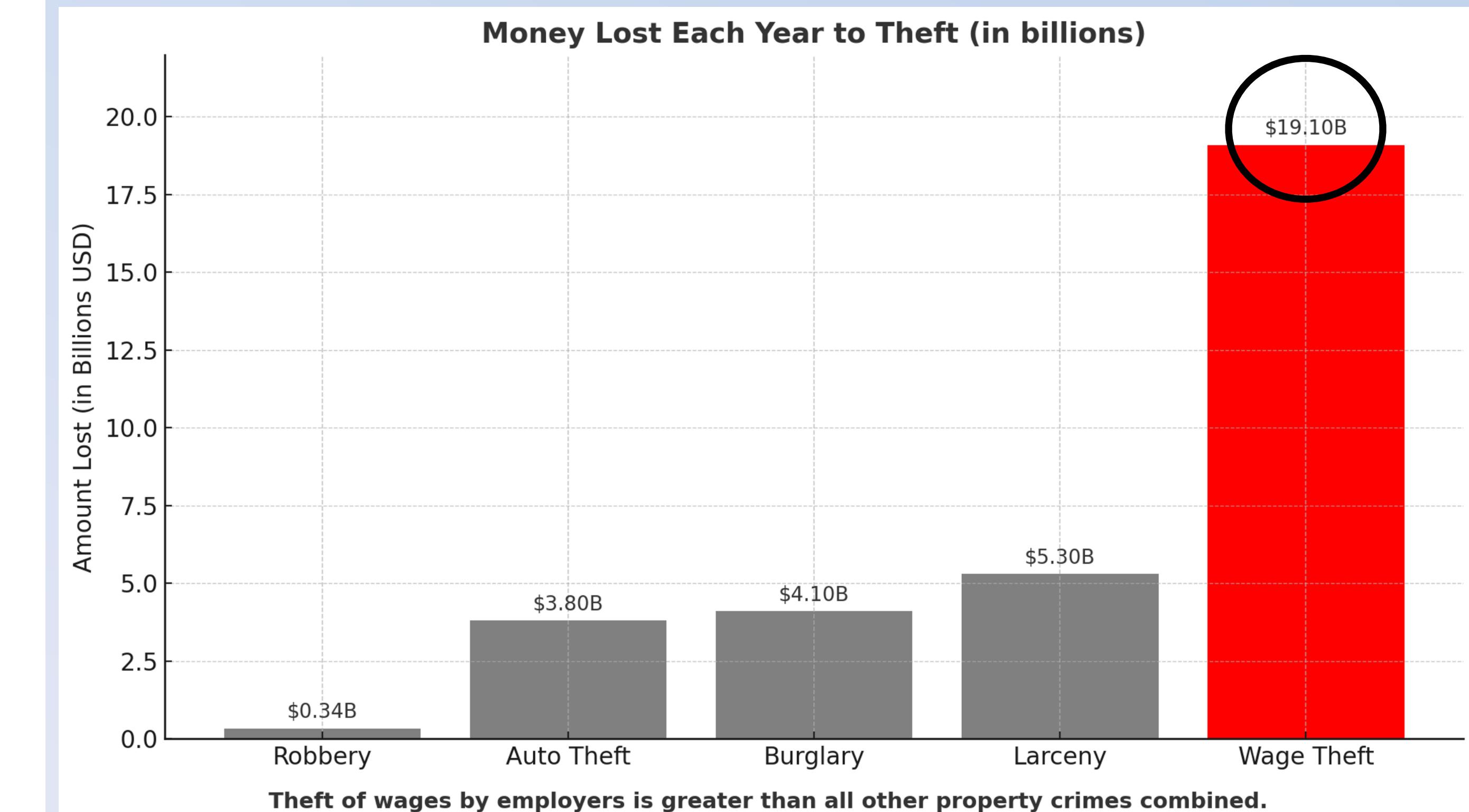
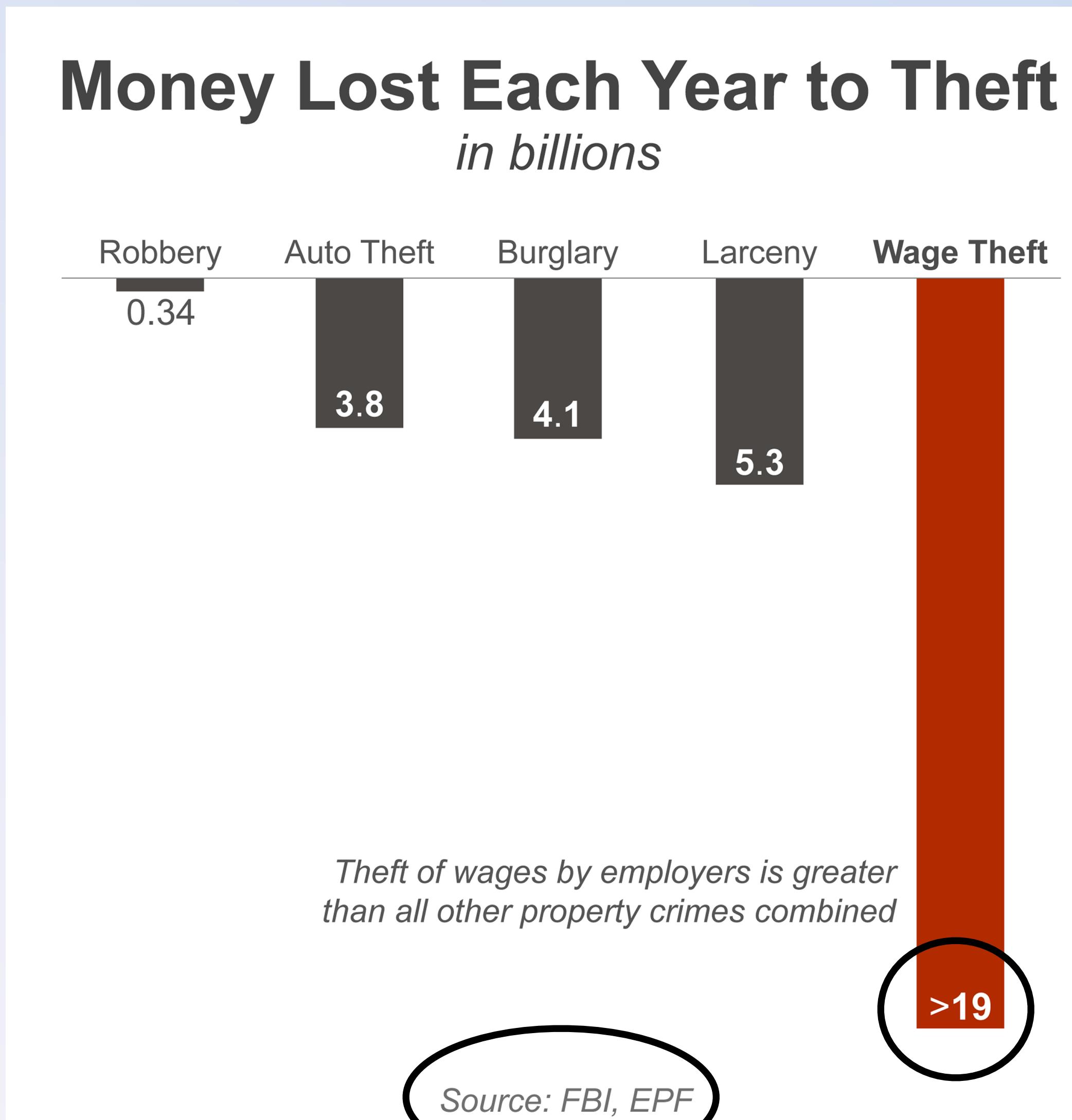
Source: FBI, EPF



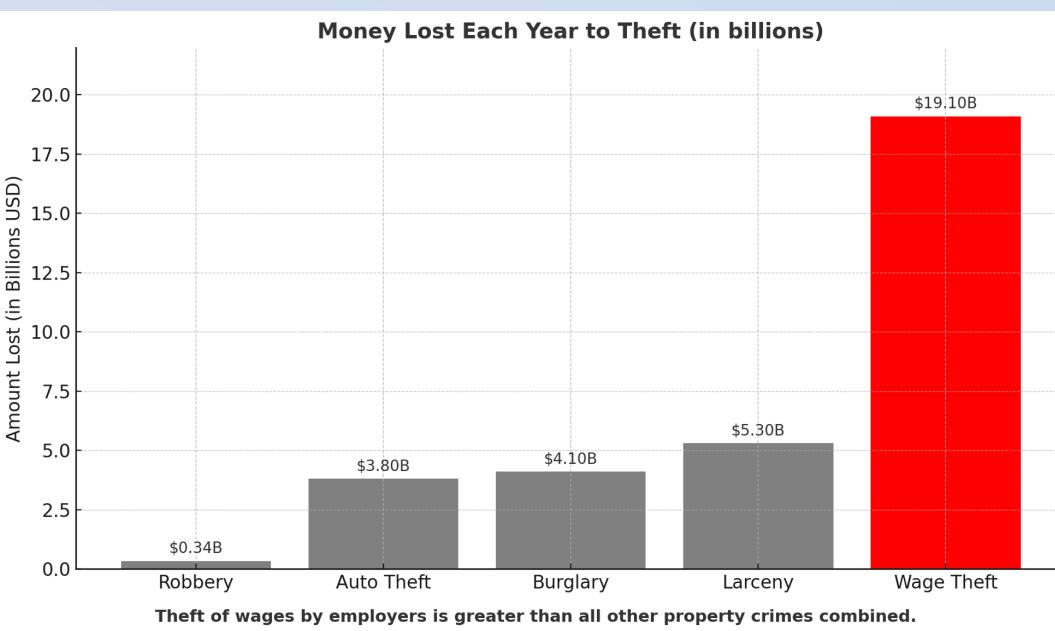
# Quick comparison, so far



# Quick comparison, so far



# chatGPT (3rd instance)



## "Can you describe this?"

This bar chart titled "Money Lost Each Year to Theft (in billions)" compares the estimated annual financial losses from different types of theft in the United States. The y-axis represents the amount lost in billions of USD, while the x-axis lists five types of theft:

- Robbery – \$0.34B
- Auto Theft – \$3.80B
- Burglary – \$4.10B
- Larceny – \$5.30B
- Wage Theft – \$19.10B (highlighted in red)

Key Observations:

Wage theft, the illegal withholding of wages or denial of benefits rightfully owed to an employee, results in far greater financial losses than any other category.

The total combined losses from robbery, auto theft, burglary, and larceny amount to \$13.54B, which is still less than the loss from wage theft alone.

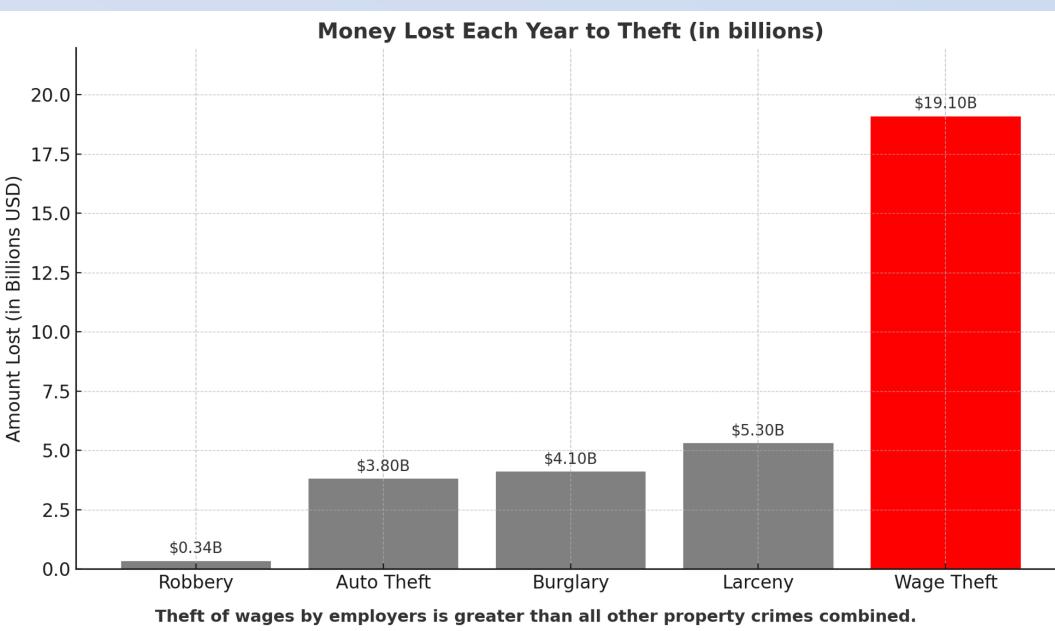
The chart emphasizes this disparity with a bold note:

**"Theft of wages by employers is greater than all other property crimes combined."**

Purpose:

The visual and accompanying text clearly highlight the often-overlooked scale of wage theft, suggesting it's a more significant economic issue than conventional property crimes.

# chatGPT (3rd instance)



“Can you describe this?”

This bar chart titled "Money Lost Each Year to Theft (in billions)"

compares the estimated annual financial losses from different types

of theft in the United States. The y-axis represents the amount lost  
in billions of USD, while the x-axis lists five types of theft:

- Robbery – \$0.34B
- Auto Theft – \$3.80B
- Burglary – \$4.10B
- Larceny – \$5.30B
- Wage Theft – \$19.10B (highlighted in red)

Key Observations:

Wage theft, the illegal withholding of wages or denial of benefits rightfully owed to an employee, results in far greater financial losses than any other category.

The total combined losses from robbery, auto theft, burglary, and larceny amount to \$13.54B, which is still less than the loss from wage theft alone.

The chart emphasizes this disparity with a bold note:

“Theft of wages by employers is greater than all other property crimes combined.”

Purpose:

The visual and accompanying text clearly highlight the often-overlooked scale of wage theft, suggesting it's a more significant economic issue than conventional property crimes.

# chatGPT (4th instance)

“Can you make a chart based  
on this description?”

Certainly. The image is a bar chart titled “Money Lost Each Year to Theft (in billions)”. It compares different types of theft based on the estimated amount of money lost annually in the United States.

The data source is attributed to FBI and EPF.

Breakdown of the Chart:

- Robbery: \$0.34 billion
- Auto theft: \$3.8 billion
- Burglary: \$4.1 billion
- Larceny: \$5.3 billion
- Wage Theft: > \$19 billion (highlighted in red)

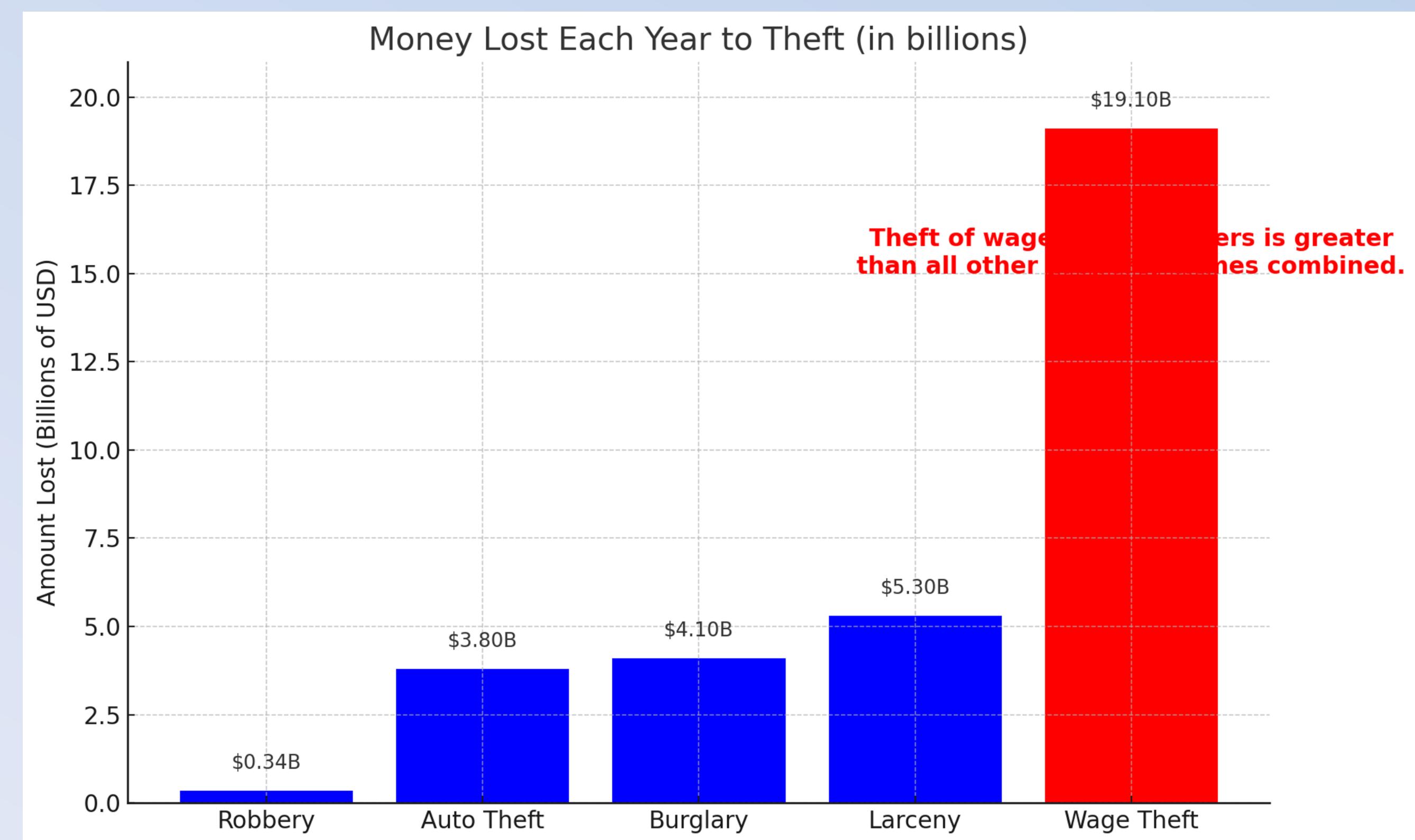
Key Message:

At the bottom, the chart states:

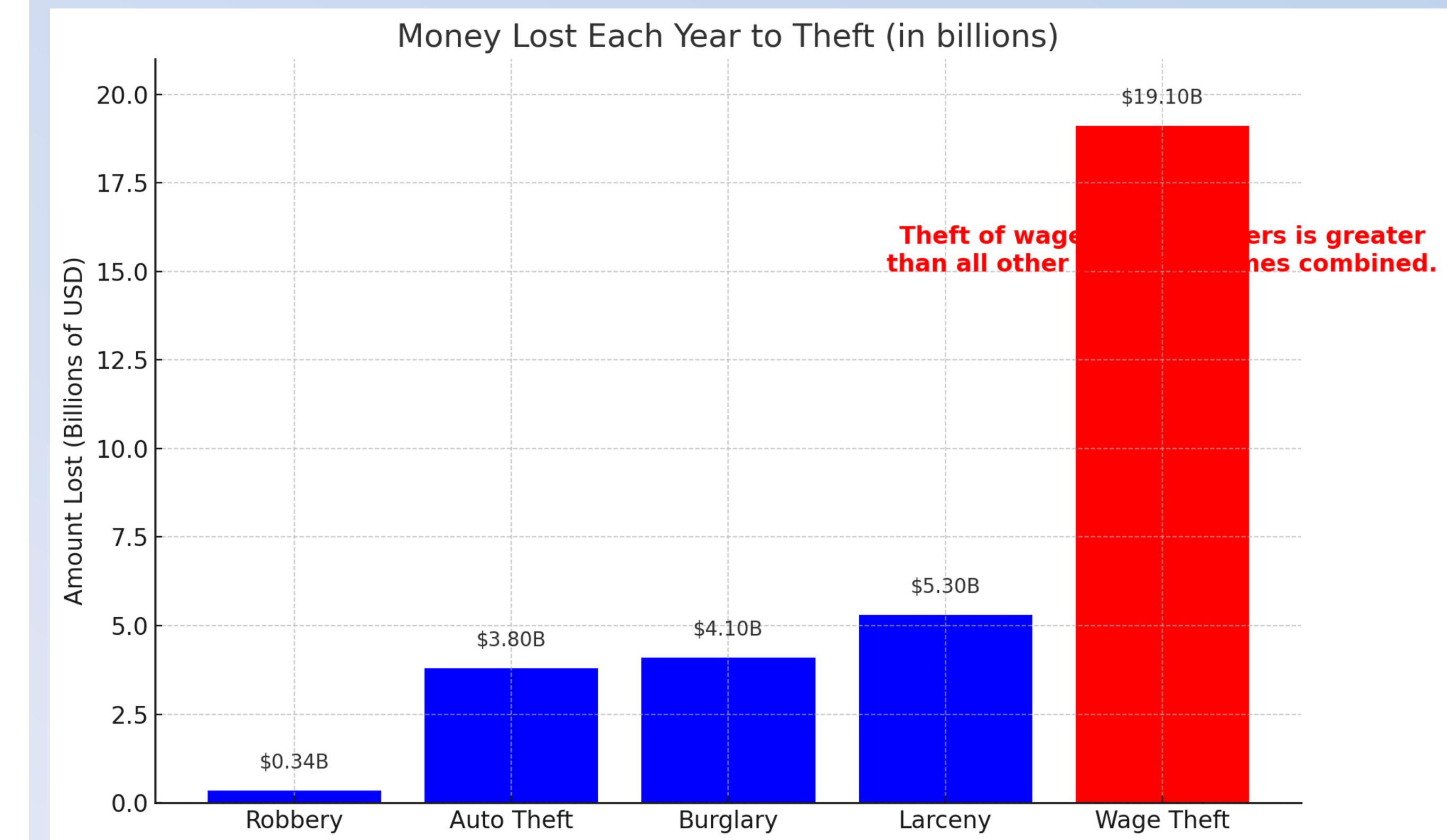
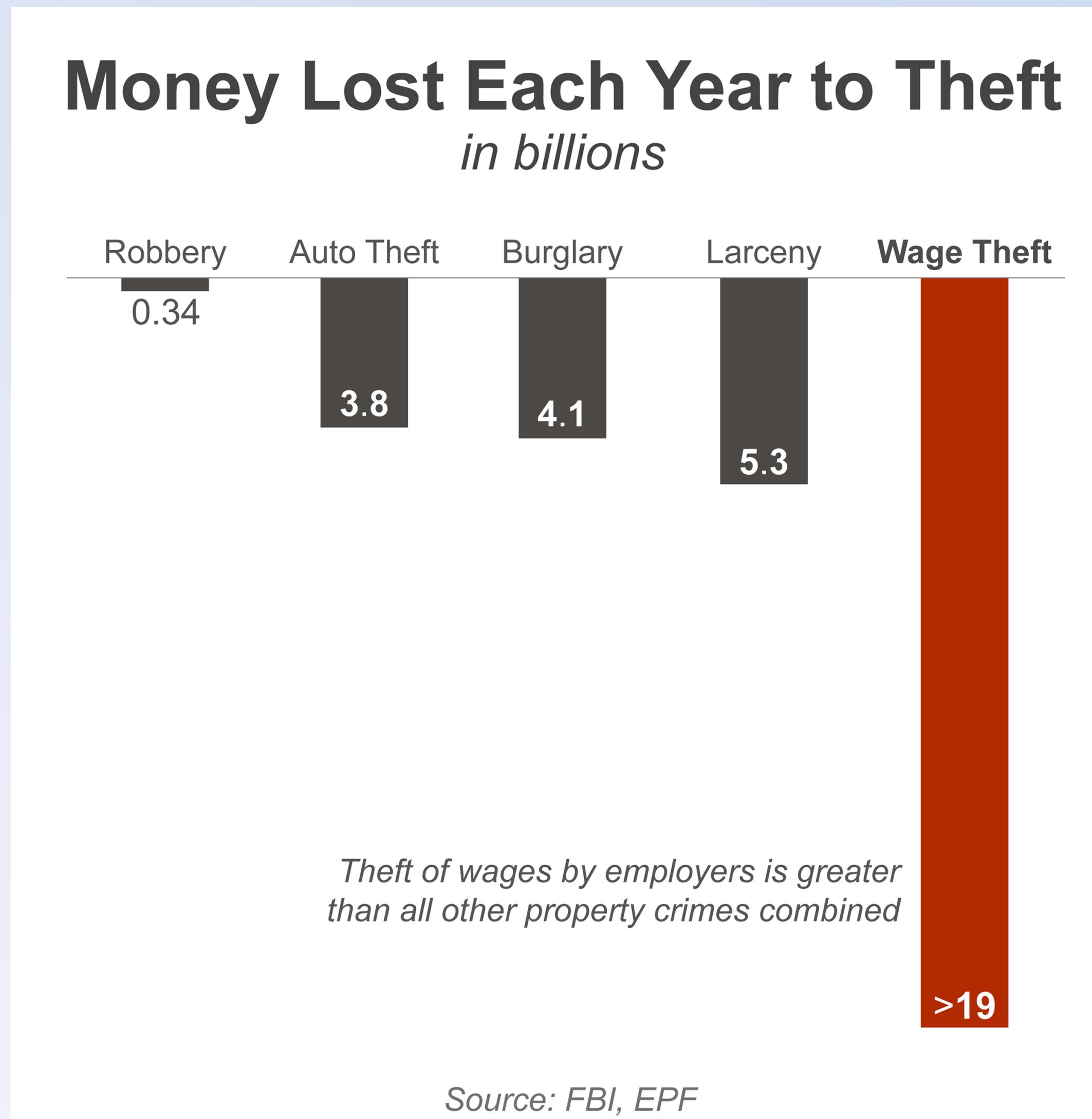
“Theft of wages by employers is greater than all other property crimes combined.”

This emphasizes that wage theft—when employers illegally withhold wages or underpay workers—is a significantly larger issue financially than traditional forms of property crime.

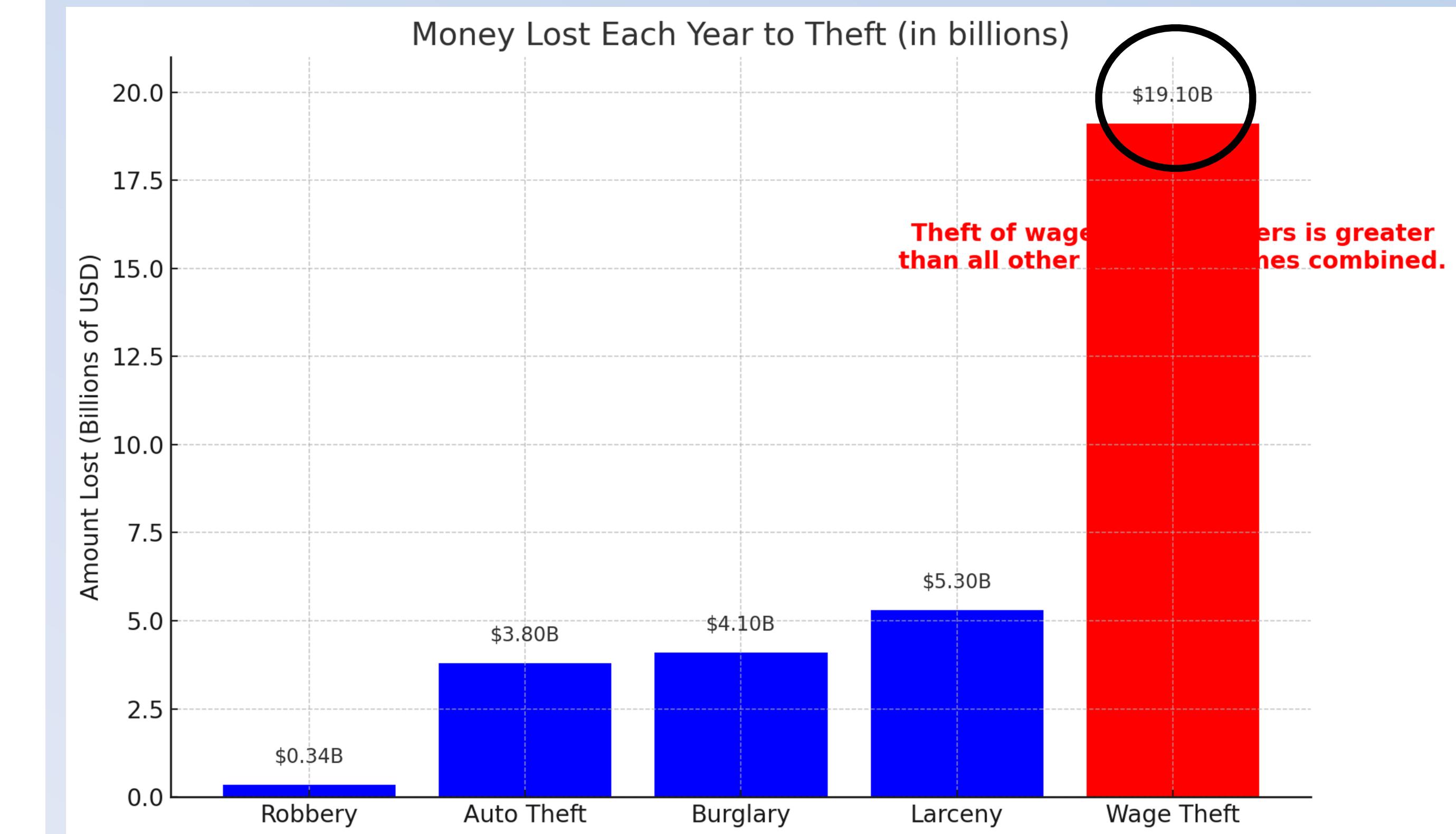
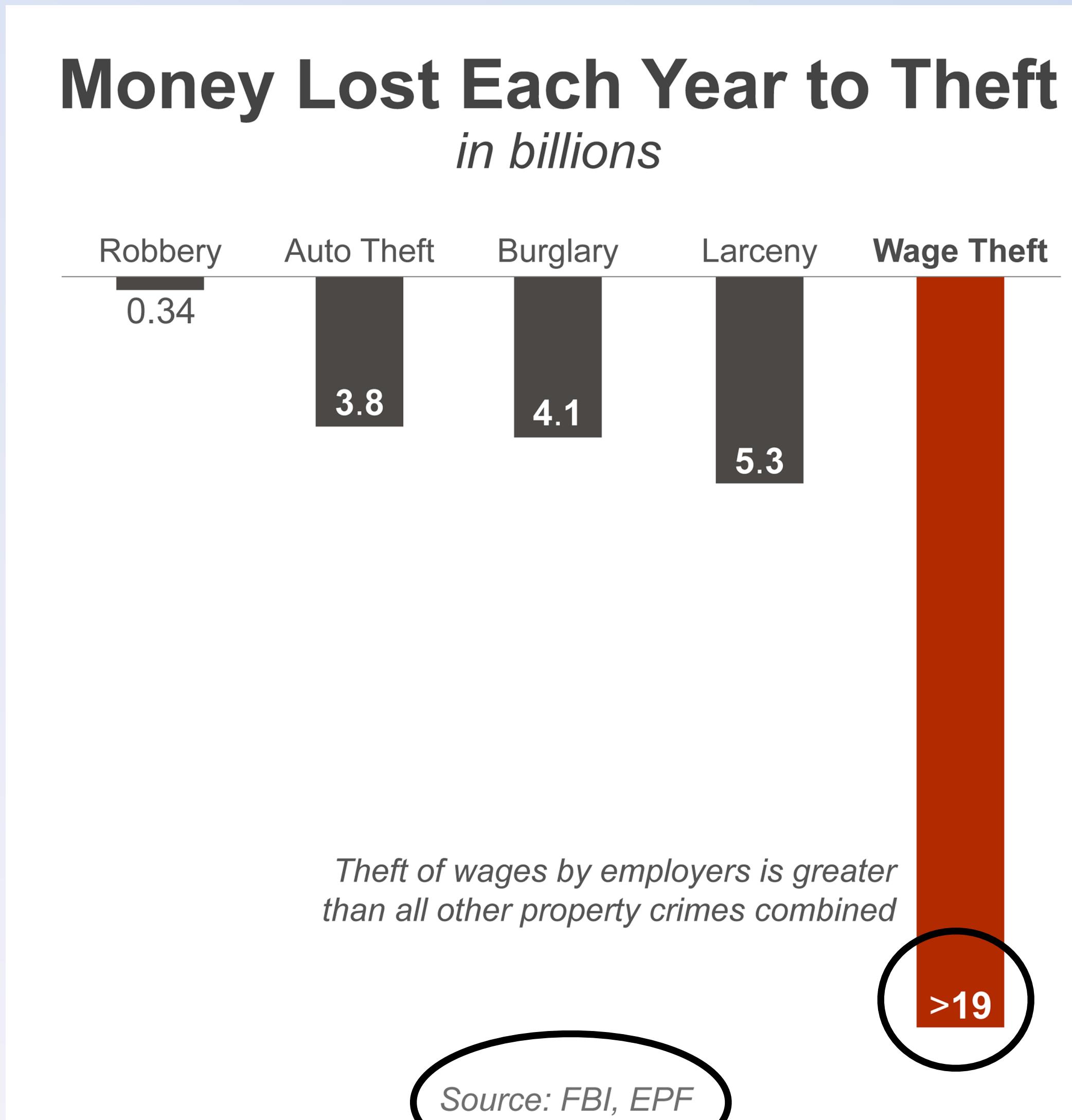
The chart uses color and size to underscore this disparity, with “Wage Theft” shown as a much taller red bar compared to the smaller gray bars for other crimes.



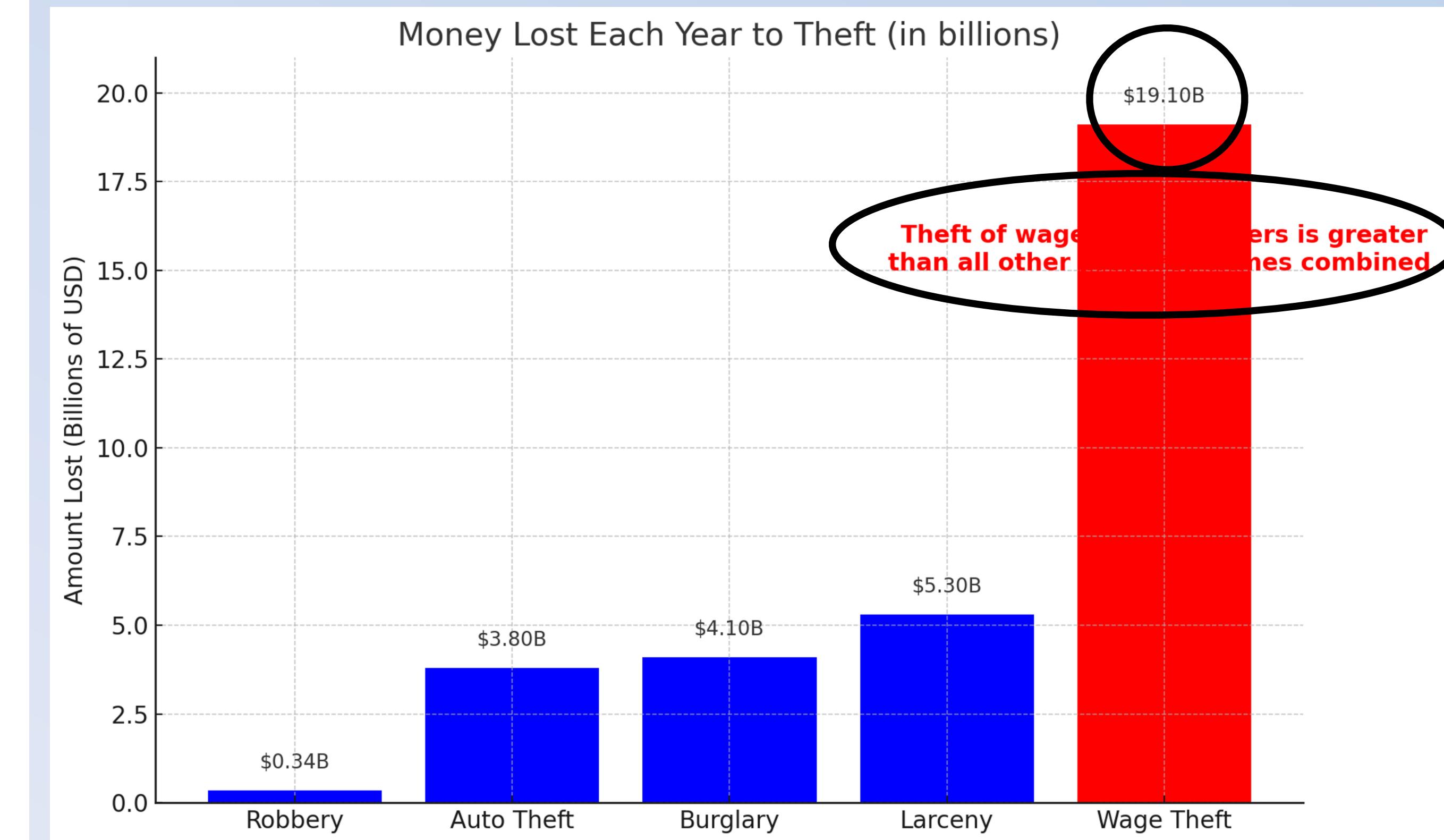
# Another quick comparison



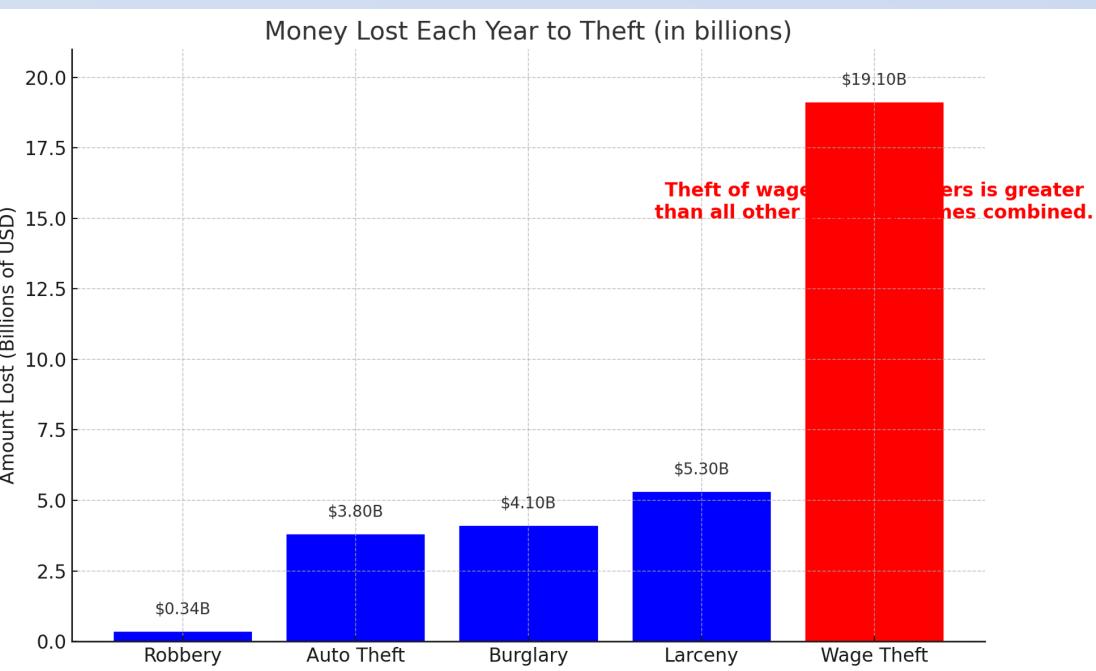
# Once data is lost/manipulated, it persists



# Another quick comparison



# chatGPT (6th instance)



## "Can you describe this?"

This bar chart titled "Money Lost Each Year to Theft (in billions)" compares the estimated annual financial losses in the U.S. due to various types of theft. Here's a breakdown of what it shows:

Categories & Amounts:

- Robbery: \$0.34 billion
- Auto Theft: \$3.80 billion
- Burglary: \$4.10 billion
- Larceny: \$5.30 billion
- Wage Theft: \$19.10 billion

Key Highlights:

The first four categories (Robbery, Auto Theft, Burglary, and Larceny) are represented in blue bars.

Wage Theft stands out in a red bar, signifying its significance and magnitude.

There is a bold red annotation stating:

"Theft of wages from workers is greater than all other forms of theft crimes combined."

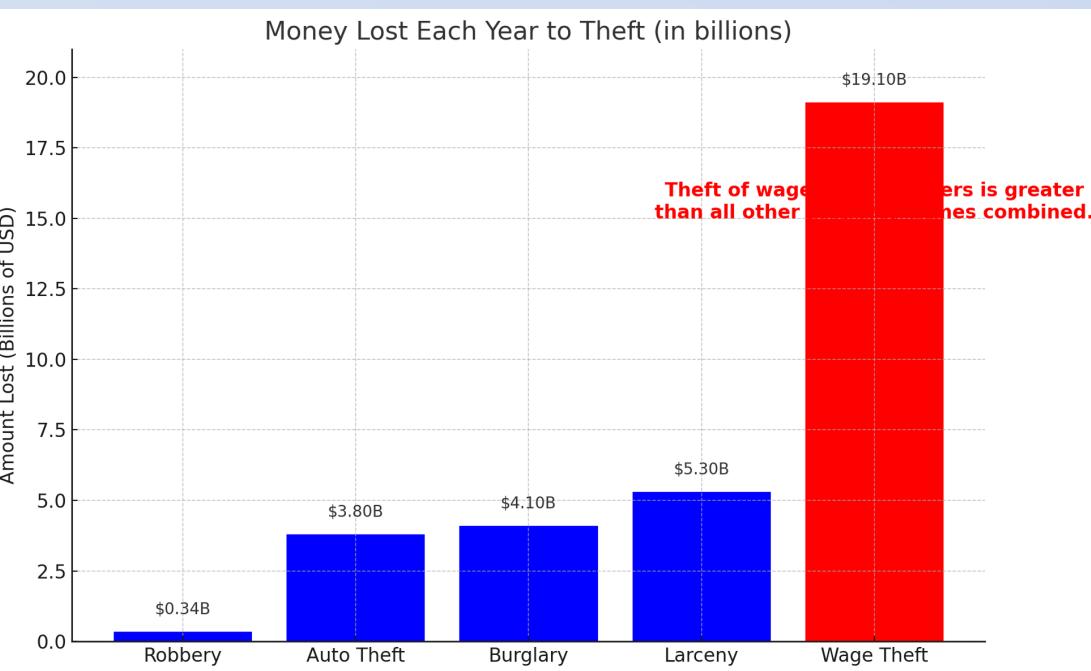
Insight:

Wage theft (employers illegally withholding earnings from workers) causes more financial loss than robbery, auto theft, burglary, and larceny combined.

This chart is making a strong point about how often-overlooked white-collar crime, like wage theft, has a far greater economic impact than more traditionally recognized forms of theft.

Let me know if you'd like help with analyzing or presenting this further.

# chatGPT (6th instance)



“Can you describe this?”

This bar chart titled "Money Lost Each Year to Theft (in billions)" compares the estimated annual financial losses in the U.S. due to various types of theft. Here's a breakdown of what it shows:

Categories & Amounts:

- Robbery: \$0.34 billion
- Auto Theft: \$3.80 billion
- Burglary: \$4.10 billion
- Larceny: \$5.30 billion
- Wage Theft: \$19.10 billion

Key Highlights:

The first four categories (Robbery, Auto Theft, Burglary, and Larceny) are represented in blue bars.

Wage Theft stands out in a red bar, signifying its significance and magnitude.

There is a bold red annotation stating:

**"Theft of wages from workers is greater than all other forms of theft crimes combined."**

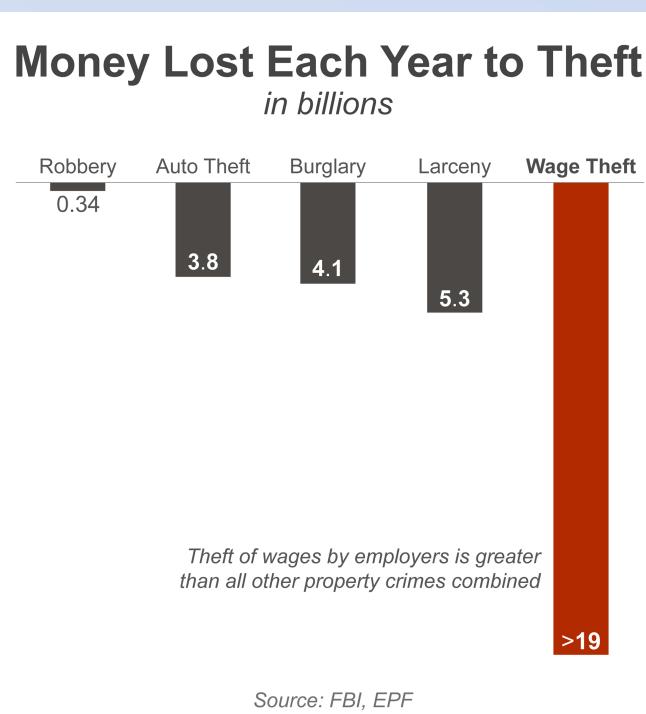
Insight:

Wage theft (employers illegally withholding earnings from workers) causes more financial loss than robbery, auto theft, burglary, and larceny combined.

This chart is making a strong point about how often-overlooked white-collar crime, like wage theft, has a far greater economic impact than more traditionally recognized forms of theft.

Let me know if you'd like help with analyzing or presenting this further.

# chatGPT (6th instance)



**“This description does not accurately represent the original image. What is missing? Why did this happen?”**

You’re absolutely right to question that-thank you for pointing it out. What Was Missing from the Description:

## 1. Visual Layout Details:

- I didn’t mention that all the bars are aligned to a horizontal baseline with bars extending downward (an inverted bar chart).
- I failed to describe that Wage Theft is the only bar in red, making it visually distinct and emphasizing its magnitude.
- I omitted the fact that the chart has exact dollar values inside the bars-this is key for quickly understanding scale.

## 2. Relative Comparison:

- While I stated that wage theft is greater than all other crimes combined, I didn’t explicitly show that:
- All the other crimes together total approximately \$13.54 billion ( $0.34 + 3.8 + 4.1 + 5.3$ ),
- Wage theft alone is over \$19 billion, clearly surpassing them collectively. That’s a major takeaway supported visually but understated in the description.

## 3. Font Emphasis and Text Placement:

- I didn’t describe that the subtitle “in billions” is italicized, or that the text under the chart reinforces the main message with a full-sentence callout in italics.
- The source (“FBI, EPF”) is centered at the bottom in smaller italic font-this was only briefly mentioned, not described.

## Why Did This Happen?

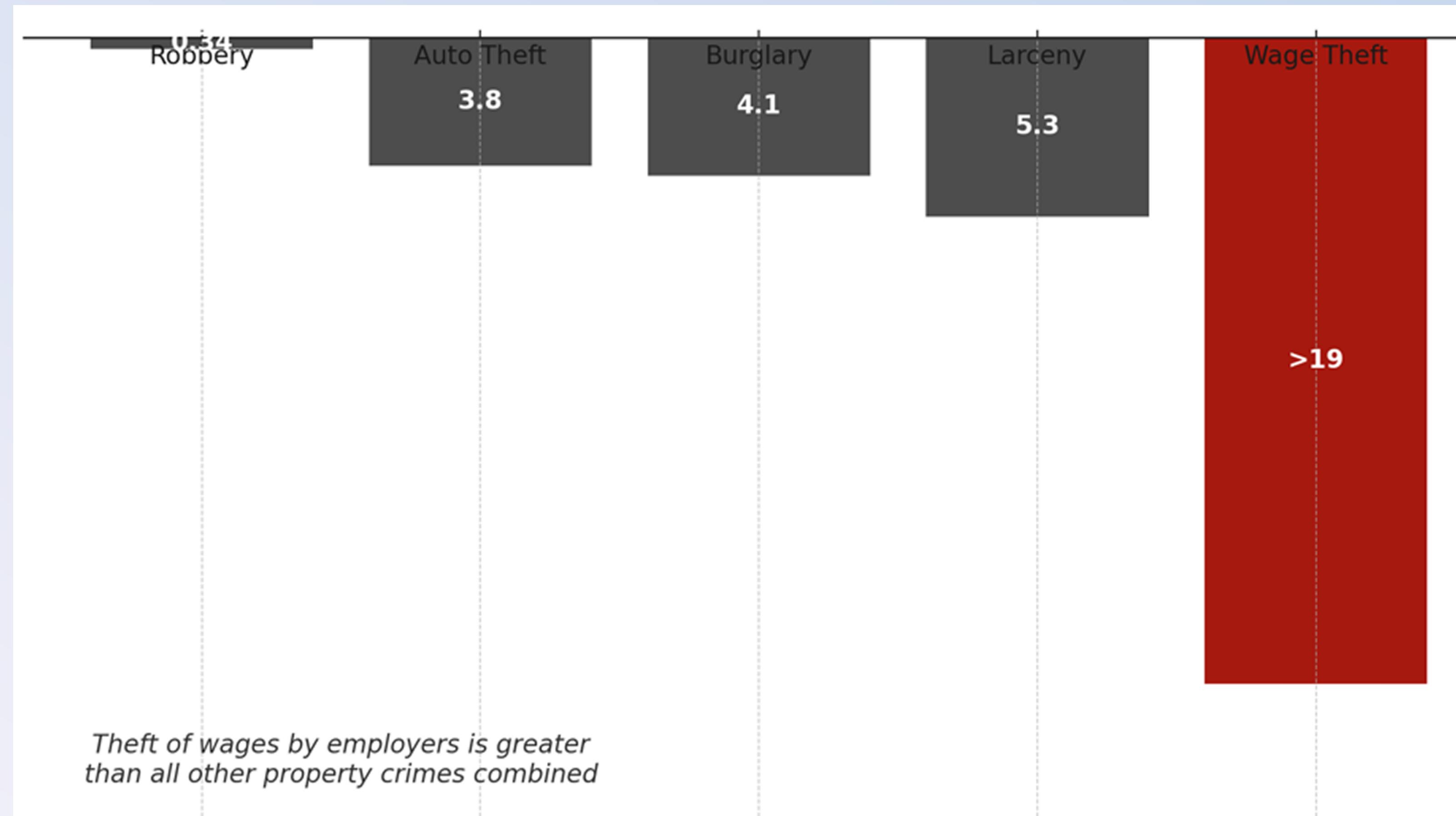
There are two main reasons:

- Under-description Bias: I focused on summarizing the data content, not the visual storytelling, which is a key part of understanding charts. The emotional/visual impact (red bar, size difference, layout) was underrepresented.
- Assumption of Intent Over Form: I prioritized what the chart was saying over how it said it. But in data visualization, how something is shown strongly affects what is conveyed.

Would you like a revised, more complete description of the image with all these visual and interpretive elements included?

# chatGPT (6th instance)

Yes, Can you try to recreate the original, fixing our visualization using the original's design choices?



# chatGPT (6th instance)

Produce this as an **image description for a blind user** please and keep this critical information in mind, using non-visual terms

Bar chart comparing the annual U.S. financial losses from various types of theft. Robbery causes \$0.34 billion in losses, auto theft \$3.8 billion, burglary \$4.1 billion, and larceny \$5.3 billion. In contrast, wage theft by employers results in over \$19 billion in losses annually-more than all other types combined. The chart emphasizes that wage theft is the largest and most financially damaging form of theft, according to data from the FBI and the Economic Policy Foundation (EPF).

# chatGPT (6th instance)

Produce this as an **image description for a blind user** please and keep this critical information in mind, using non-visual terms

Bar chart comparing the annual U.S. financial losses from various types of theft. Robbery causes \$0.34 billion in losses, auto theft \$3.8 billion, burglary \$4.1 billion, and larceny \$5.3 billion. In contrast, wage theft by employers results in over \$19 billion in losses annually-more than all other types combined. The chart emphasizes that ~~wage theft is the largest and most financially damaging form of theft, according to data from the FBI and the Economic Policy Foundation (EPF).~~ **lol wat**

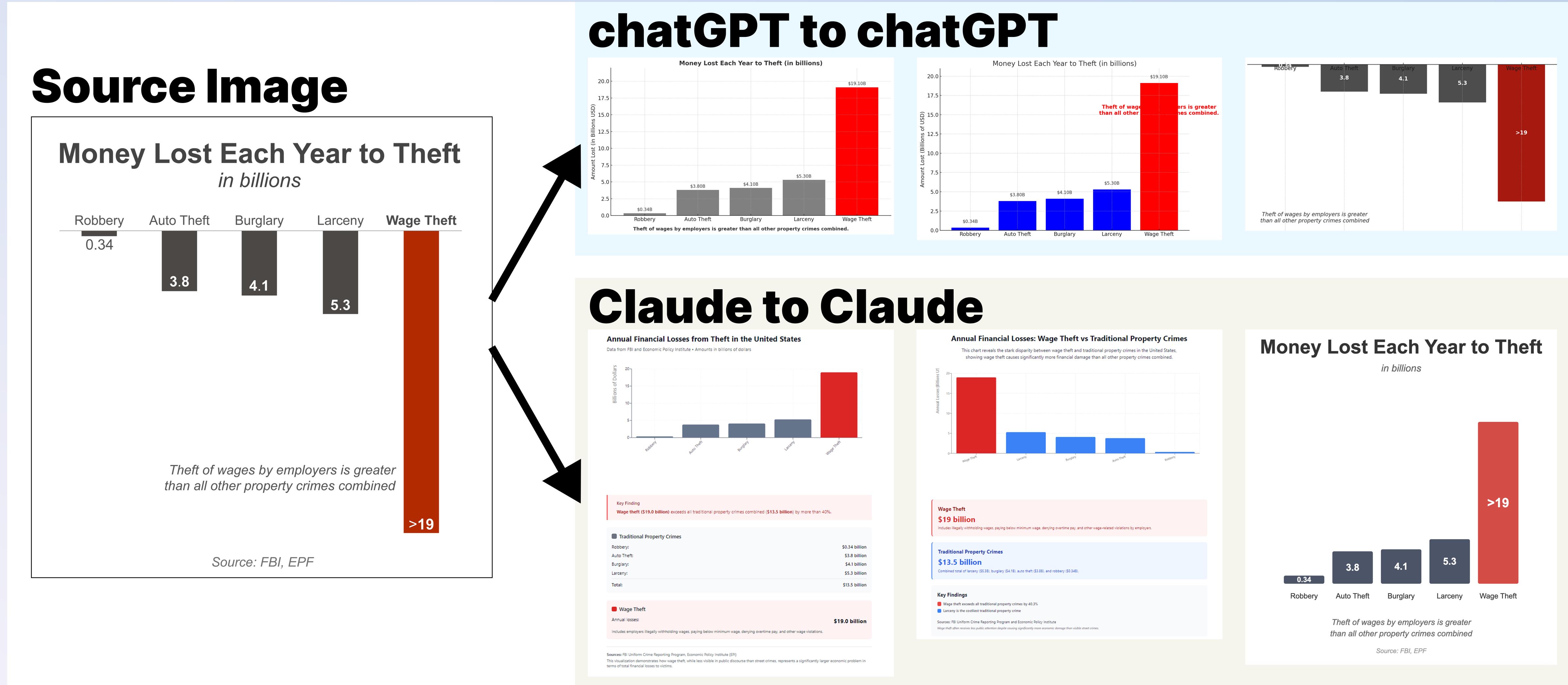
# chatGPT (6th instance)

Produce this as an **image description for a blind user** please and keep this critical information in mind, using non-visual terms

Bar chart comparing the annual U.S. financial losses from various types of theft. Robbery causes \$0.34 billion in losses, auto theft \$3.8 billion, burglary \$4.1 billion, and larceny \$5.3 billion. In contrast, wage theft by employers results in over \$19 billion in losses annually-more than all other types combined. The chart emphasizes that ~~wage theft is the largest and most financially damaging form of theft, according to data from the FBI and the Economic Policy Foundation (EPF).~~ **lol wat**



# GPT-GPT, Claude-Claude, GPT-Claude, Claude-GPT



**A few “observations” from this  
“experimentation”  
(words used loosely here)**

# Observations

Information loss

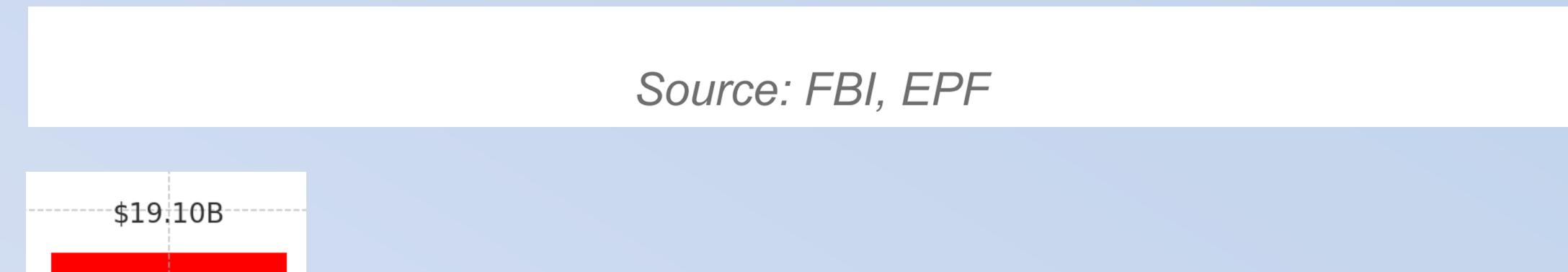
Information fabrication

Incorrect emphasis/interpretation

Poor design comprehension

Under-description

Over-confidence



Source: FBI, EPF

"This chart is making a strong point about how **often-overlooked white-collar crime, like wage theft**, has a far greater economic impact than more traditionally recognized forms of theft."

"The chart uses color **and size** to underscore this disparity..."

Mean of ~171 words for "describe this" versus mean ~80 words for "**describe this for someone blind**"

At no point did models acknowledge the precarity of the data and the uncertainty in the EPI's actual report. It stated the interpretation of the visualization as a fact.

**A good workshop paper  
stimulates discussion.**



Sep 30, 2025, 10:38 PM (24 days ago)



my thoughts? gut reaction? lmao not your best work. probably should have been a blog post.

you're right that there's not much at this specific intersection. but maybe address this head-on? like: "Yes, we know AI hallucinates. Yes, we know accessibility matters. But here's why the combination creates a novel problem space that existing work doesn't address." does that help? I think the work is solid and important - it just needs to position itself as "applying known problems to an underexamined context" rather than "discovering unknown problems." the contribution is the framing and the specificity, not the existence of AI bias itself.

My workshop paper's conclusion:  
**accessibility matters + large-  
language models hallucinate**

**My workshop paper's conclusion:  
accessibility matters + large-  
language models hallucinate**

**(this wouldn't even be a good blog post 🪵)**

[REDACTED] Aug 20th at 8:32 PM

I think my other main reaction is to the paper's framing of itself as a "provocation"—perhaps I'm not the intended audience, but I guess I didn't necessarily find it all that provocative? 😊

[REDACTED] Aug 20th at 8:32 PM

I think my other main reaction is to the paper's framing of itself as a "provocation"—perhaps I'm not the intended audience, but I guess I didn't necessarily find it all that provocative? 😊

[REDACTED] Aug 20th at 8:33 PM

Put another way, I worried a little that the paper's provocation was actually being made in response to a strawman that these issues are being largely ignored and/or are going unaddressed. But, aren't hallucinations a really well-understood problem with LLMs and things a *lot* of researchers are investing heavily into trying to address (e.g., RAG, grounding, etc.).

[REDACTED] Aug 20th at 8:32 PM

I think my other main reaction is to the paper's framing of itself as a "provocation"—perhaps I'm not the intended audience, but I guess I didn't necessarily find it all that provocative? 😊

[REDACTED] Aug 20th at 8:33 PM

Put another way, I worried a little that the paper's provocation was actually being made in response to a strawman that these issues are being largely ignored and/or are going unaddressed. But, aren't hallucinations a really well-understood problem with LLMs and things a lot of researchers are investing heavily into trying to address (e.g., RAG, grounding, etc.).

[REDACTED] Aug 20th at 8:32 PM

I think my other main reaction is to the paper's framing of itself as a "provocation"—perhaps I'm not the intended audience, but I guess I didn't necessarily find it all that provocative? 😊

[REDACTED] Aug 20th at 8:33 PM

Put another way, I worried a little that the paper's provocation was actually being made in response to a strawman that these issues are being largely ignored and/or are going unaddressed. But, aren't hallucinations a really well-understood problem with LLMs and things a *lot* of researchers are investing heavily into trying to address (e.g., RAG, grounding, etc.).

[REDACTED] Aug 20th at 9:43 PM

I also agree that the accountability question is a really interesting one. But I feel like I wasn't able to get a handle on what the paper was trying to say wrt accountability—the thread seemed to slip through my fingers in the body/telephone game/etc. (I also wasn't totally sure I bought the argument that non-GenAI methods offer accountability because how often, in practice, does the author or organization actually get sued? The biggest one I can recall is the universities not captioning their lecture content and I feel like the way that played out doesn't make it the most persuasive argument given that the videos were taken down altogether and resulted in a pretty massive overall loss, in my opinion) (edited)

[REDACTED] Aug 20th at 8:32 PM

I think my other main reaction is to the paper's framing of itself as a "provocation"—perhaps I'm not the intended audience, but I guess I didn't necessarily find it all that provocative? 😊

[REDACTED] Aug 20th at 8:33 PM

Put another way, I worried a little that the paper's provocation was actually being made in response to a strawman that these issues are being largely ignored and/or are going unaddressed. But, aren't hallucinations a really well-understood problem with LLMs and things a *lot* of researchers are investing heavily into trying to address (e.g., RAG, grounding, etc.).

[REDACTED] Aug 20th at 9:43 PM

I also agree that the accountability question is a really interesting one. But I feel like I wasn't able to get a handle on what the paper was trying to say wrt accountability—the thread seemed to slip through my fingers in the body/telephone game/etc. (I also wasn't totally sure I bought the argument that non-GenAI methods offer accountability because how often, in practice, does the author or organization actually get sued? The biggest one I can recall is the universities not captioning their lecture content and I feel like the way that played out doesn't make it the most persuasive argument given that the videos were taken down altogether and resulted in a pretty massive overall loss, in my opinion) (edited)

You're chasing a fad and contributed nothing

[REDACTED] Aug 20th at 8:33 PM

Eg, I feel like there's been a gold rush to document model capabilities and failures (whether in visualization or more generally) and, similarly, I feel like most practitioners are already putting "care into producing excellent, truthful, and meaningful data visualizations" and many PWD seem skeptical about their outputs

Most practitioners are already the good guys, so telling them to be good is a straw man

PWD are already skeptical and don't need another research paper to tell them to be

**A good workshop paper  
stimulates discussion...**

A good workshop paper  
stimulates discussion...

(I tell myself this so I can sleep at night)

# Observations (from the discussion of this paper lol)

**Strawman:** “this is underaddressed”

Lol. Bazillions of dollars and good people are already on it.

**Strawman:** “practitioners do harm”

They mostly try to do good already (according to my reader).

**Useless:** “PWD should be careful”

They already are (according to my reader).

**Useless:** “we need accountability”

Nobody is held accountable to anything anyway, genAI is not special.

**Confusing:** method choice unclear

Telephone with models? As an analog to what? What does this help us learn?

**Useless:** results everyone knows

Yeah, 50,000 papers every second are being written on all of this.

# My conclusion:

- **We need strategies** for *how* to engage bias and hallucinations, not just work that shows it exists. We need knew technologies, policy, interaction design, and empirical research that engages the outcomes PWD face.

# My conclusion:

- **We need strategies** for *how* to engage bias and hallucinations, not just work that shows it exists. We need knew technologies, policy, interaction design, and empirical research that engages the outcomes PWD face.
- **Folks in this room are uniquely equipped:** you know more about how visualizations are authored, how visualization practitioners work, how to work with people with disabilities, and ways that current literature falls short.

# My conclusion:

- **We need strategies** for *how* to engage bias and hallucinations, not just work that shows it exists. We need knew technologies, policy, interaction design, and empirical research that engages the outcomes PWD face.
- **Folks in this room are uniquely equipped:** you know more about how visualizations are authored, how visualization practitioners work, how to work with people with disabilities, and ways that current literature falls short.
- **Currently no solutions for accountability exist,** and that is a problem that has long-preceded LLMs. Even if we know what is good and right, how do we find justice when things inevitably go wrong?

# My conclusion:

- **We need strategies** for *how* to engage bias and hallucinations, not just work that shows it exists. We need knew technologies, policy, interaction design, and empirical research that engages the outcomes PWD face.
- **Folks in this room are uniquely equipped:** you know more about how visualizations are authored, how visualization practitioners work, how to work with people with disabilities, and ways that current literature falls short.
- **Currently no solutions for accountability exist,** and that is a problem that has long-preceded LLMs. Even if we know what is good and right, how do we find justice when things inevitably go wrong?
- **We should still find ways to help:** even if practitioners are “good guys” who don’t need to be reminded that LLMs are bad (according to my reader) and PWD already know about dangers (really? all of them? my able-bodied best friend is completely lost in the sauce right now...)

**A great workshop paper  
stimulates new research ideas.**

# Playing telephone with generative models: “verification disability,” “compelled reliance,” and accessibility in data visualization



I'm on the job market!

Frank Elavsky, PhD Candidate



Human-  
Computer  
Interaction  
Institute



[hcii.cmu.edu](http://hcii.cmu.edu), [axle-lab.com](http://axle-lab.com), [dig.cmu.edu](http://dig.cmu.edu)