

# How to Price AI Agents

## The 2026 Guide to Agentic AI Pricing

A comprehensive framework for pricing the products replacing the old guard. From copilots to autonomous agents, from per-seat to outcome-based models.

### FEATURING CASE STUDIES

- Cursor
- Devin
- Harvey AI
- 11x (Alice)
- Sierra AI



*SaaS and AI Pricing strategy that consistently increases revenue by 12-40%.  
Delivered via tech-enabled consulting in record speed.*

April 2026 | [getmonetizely.com](https://getmonetizely.com)

# Contents

---

## **Part 0 SaaS Is Giving Way To Agentic AI**

- The Moment Everything Changed
- A Deflation Bomb on a Category That Was Already Deflating
- The Inference Cost Collapse That Lit the Fuse
- Where the Money Is Going
- Why This Matters for Pricing
- The Five Companies That Will Teach You Agentic AI Pricing

## **Part I The 5-Step Pricing Transformation Framework**

## **Part II Packaging, or Designing Offers That Map to Your Buyer**

- Cursor: Segments Understood, Capabilities Mapped Well
- Devin: Right Segments, But the Packages Do Not Match
- Harvey AI and Sierra AI: Right Capabilities for One Segment
- 11x (Alice): One Package That Speaks to No Segment's Job

## **Part III Pricing Metric and Price Point - Enter the AMS**

- The Three Dimensions
- The Five Agents on the AMS
- How the AMS Determines Pricing Metric
- How the AMS Informs Price Point

## **Part IV What We Have Learned**

# The Moment Everything Changed

On February 3rd, 2026, about \$285 billion in market cap vanished from software stocks in 48 hours.

Anthropic launched Claude Cowork with industry plugins for legal, sales, financial analysis, and marketing. This repriced the entire SaaS segment in the market.

AI was eating SaaS at exponential pace.

## KEY INSIGHT

*"If 10 AI agents can do the work of 100 reps, you need 10 Salesforce seats, not 100."*

- Jason Lemkin

Jefferies traders coined a term for the selloff: the SaaSocalypse. Atlassian dropped 35% and reported its first ever decline in enterprise seat counts. Salesforce fell 33%. Snowflake shed 37%. Adobe fell 36%. The average forward earnings multiple for software companies went from roughly 39x to 21x in a matter of weeks.

The market reclassified the entire software sector in real time.

---



## A Deflation Bomb on a Category That Was Already Deflating

Here is something most people in software do not understand, and it changes the whole picture once you do.

Software is one of the highest deflating product categories in the economy. While eggs inflate, while tuition inflates, while medical costs inflate, software deflates. If you go to the Bureau of Labor Statistics and plot it yourself, software that was worth \$100 in 2015 is worth about \$60 today. Televisions deflate. Toys deflate. But software deflates right at that level or faster.

This was already true before AI.

AI has been an exponential accelerant. On a category that was already highly deflating, we have now set off a deflation bomb.

Y Combinator reports that 25% of their current batch has codebases that are 95% or more AI generated. Things that used to take months to develop are being developed within weekends now.

Lovable went from zero to \$100 million ARR in eight months, then doubled to \$200 million in four more. There are already 10,000 MarTech tools, 10,000 sales tech tools, and it takes a weekend to build the software anymore. When a two-person startup can ship in weeks what used to take a 50-person team 18 months, the pricing power of the incumbent evaporates.

Software is getting commoditized at a speed that still feels like a singularity event without people fully recognizing it. The deflation is hitting SaaS companies from two directions.

From below: AI coding tools make it cheap to build alternatives.

From above: AI agents make it possible to skip the software entirely and just do the work. A company that used to buy a \$50,000/year project management tool can now point an AI agent at the work and have it managed. The tool becomes optional. The work is what matters.



## The Inference Cost Collapse That Lit the Fuse

None of this would be happening if AI itself were still expensive to run.

Meeker's report documented a 99.7% drop in inference costs over two years, citing Stanford research. That reduction is even faster than Moore's Law. Nvidia's 2024 Blackwell GPU chip uses 105,000 times less energy per token than the 2014 Kepler model. Meeker called it a staggering leap, not just of cost reduction, but of architectural and materials innovation. Tunguz showed that Google's Gemini 3 Flash underprices other top models by 70 to 79% while matching their performance across 20 benchmarks.

This cost collapse is what makes the agentic economy inevitable. When inference was expensive, AI agents were demos. Interesting to watch, too costly to run at scale. When inference drops to pennies, every business process becomes a candidate for agent

automation. The unit economics flip from "can we afford to use AI for this?" to "can we afford not to?"

And the decline is accelerating. The floor for inference cost is nowhere in sight.



## Where the Money Is Going

In Q1 2026, four of the five largest venture rounds ever recorded closed: OpenAI at \$122 billion, Anthropic at \$30 billion, xAI at \$20 billion, Waymo at \$16 billion. Together, those four deals made up 65% of all global venture investment for the quarter. AI companies absorbed 80% of total global venture funding in Q1 2026, up from 55% just a year prior. Sequoia's annual AI report found that \$28 billion in venture capital flowed into AI agent startups in 2025 alone, a 4x increase from 2024.

Y Combinator's Spring 2025 batch was 46% AI agents. Nearly half.

And this goes beyond VC hype. Tunguz of Theory Ventures predicted 2026 would be the year agents out-earn people. It is already happening. Tunguz found that Asana's internal data showed a statistically significant "Agent Inflection Point" on October 12, 2025, where daily automated task volume tripled and shifted toward higher-value work.

Sequoia Capital framed this in their March 2026 piece "Services: The New Software." Their argument: the era of the copilot (AI assists a human using software) is giving way to the era of the autopilot (AI does the work the software was designed to support). A copilot sells the tool. An autopilot sells the work. When the autopilot arrives, the tool layer collapses. You do not need a CRM if an agent runs your pipeline. You do not need a ticketing system if an agent resolves your tickets.

The SaaS economy as we have known it for twenty years is over. The per-seat subscription model that built Salesforce, Workday, ServiceNow, Atlassian, and hundreds of others is being replaced by a new commercial architecture where the unit of value is not the human using the software but the agent doing the work.

The old world had low interest rates, manageable deflation, and high-margin software that could sustain itself indefinitely. That world is gone. Software is deflating. Inference is

deflating. Margins are compressing. The only thing that is not deflating is the value of the work that agents can do and the services wrapped around deploying them.

The capital has moved. The budgets have followed.

The SaaS era is behind us. The Agentic AI economy is here.

The question is no longer whether this shift is happening. The question is how you price for it.



## Why This Matters for Pricing

The headlines focus on destruction. Which SaaS companies will die. Which stocks will crater. But destruction is only half the story. Someone has to build the pricing models for the thing that comes next.

An AI agent that costs \$0.50 in compute to produce a \$3,200 legal memo. An autonomous SDR that sends 10,000 emails to book 15 meetings. A customer service agent that resolves tickets for pennies that used to cost \$5 to \$15 each. The value gap between what these agents produce and what they cost to run has no precedent in the software industry. And the competitive dynamics (50+ AI SDR startups in 18 months, inference costs dropping 10x per year, open-source alternatives closing the gap) mean that capturing that value gap is harder than it looks.

Get the pricing right, and you build a business that scales with the value it creates. Get it wrong, and you either leave billions on the table or price yourself out of a market moving faster than any in software history.

That is what this book is about. How to price the products that are replacing the old guard. How to build commercial models for software that thinks, acts, and delivers outcomes on its own.

We start with the 5-Step Pricing Transformation Framework that we use at Monetizely for any software product, whether SaaS, AI, or otherwise. Then we apply it to agentic AI using a framework we developed called the Agentic Monetization Spectrum (AMS), which classifies agents and maps them to the right pricing decisions.

If you sell an AI agent, chances are you have already had the pricing conversation that goes nowhere. Someone on your team says "let's do outcome-based pricing." Someone else says "per seat is safer." The conversation turns into a debate about what Devin or Harvey or Sierra is doing, as if copying another company's pricing page counts as a strategy. It does not. Pricing an AI agent is a new context applied to a set of decisions that people have been making for decades.

The physics of pricing (who you sell to, what you package, what metric you charge on, and what price you set) have not changed because the product runs on an LLM instead of a database. What has changed is the commercial properties of the product, and those properties create new tensions within the same old framework.

# The Five Companies That Will Teach You Agentic AI Pricing

---

Every concept in this book weaves through five real AI agent companies. If you understand how pricing works and fails for these five, you will have the pattern recognition to price any AI agent.

## Cursor

[AI Code Editor](#) | [\\$29.3B valuation](#) | [~\\$2B ARR](#)

An AI code editor built by Anysphere, founded in 2022 by four MIT students who forked Visual Studio Code and built AI into the entire development workflow. Cursor helps developers generate, refactor, and edit code using full codebase context, going well beyond autocomplete into multi-file, agent-driven edits.

The company crossed \$100M in ARR in January 2025, passed \$500M by June, cleared \$1B by November, and reportedly sits near \$2B ARR as of early 2026. All with zero marketing spend. Over a million daily users including teams at OpenAI, Uber, Spotify, and Instacart.

We include Cursor because it is the most proven form of AI agent commercially: the human-in-the-loop copilot. The developer is still the one doing the work. This makes it the closest thing to traditional SaaS in our set, and the pricing reflects it: per seat, tiered, familiar. The question Cursor forces us to ask is whether traditional SaaS pricing can hold up when the agent's capability keeps growing faster than the price.

---

## Devin

[Autonomous AI Software Engineer](#) | [\\$10.2B valuation](#) | [~\\$150M+ ARR](#)

An autonomous AI software engineer built by Cognition AI, founded in 2023 by a team of competitive programming champions with multiple International Olympiad in Informatics gold medals. Unlike Cursor, Devin does not assist a developer. It is the developer. You give it a task in plain language, and Devin plans the approach, writes the code, tests it, debugs it, and submits a pull request.

Cognition's ARR grew from \$1 million in September 2024 to \$73 million by June 2025, and after buying Windsurf, combined ARR is estimated above \$150 million.

We include Devin because it shows the leap from copilot to autonomous agent. Same domain as Cursor (software engineering), but a very different product to sell. When the agent does the work, per-seat pricing breaks down. Cognition addressed this with the ACU (Agent Compute Unit) model.

---

## Harvey AI

[AI Platform for Legal Work | \\$11B valuation | \\$190M ARR](#)

An AI platform for legal work: contract analysis, M&A due diligence, regulatory compliance, multi-jurisdictional research, litigation support, and memo drafting. Founded in 2022 by Winston Weinberg, a former securities litigator at O'Melveny & Myers, and Gabriel Pereyra, a former research scientist at DeepMind.

Designed for Am Law 100 firms and Fortune 500 legal departments. More than 25,000 custom agents now run on the platform.

We include Harvey because it is the high-value, domain expert agent. A legal memo that takes a junior associate 6 hours and \$3,000+ in billable time might consume \$0.50 to \$2.00 in inference. That ratio points clearly to value-based pricing. And yet Harvey uses per-seat pricing at roughly \$1,200 per lawyer per month.

---

## 11x (Alice)

[AI SDR Agent | Most competitive AI agent category](#)

Alice prospects, researches leads, personalizes outreach across email and LinkedIn, manages follow-ups, and books meetings, all running 24/7. The AI SDR space went from a handful of players to 50+ in under 18 months.

Competitors range from Agent Frank at \$499/month to AiSDR at \$900/month to 11x at about \$5,000/month.

We include 11x because it sits at the premium end of a market racing to the bottom, and because its pricing (flat monthly fee, no usage or outcome component) represents the most common and also the most exposed approach in agentic AI.

---

## Sierra AI

### Enterprise AI Customer Service | \$10B valuation

Enterprise AI agents for customer service. Founded by Bret Taylor (former Salesforce co-CEO) and Clay Bavor (former Google VP). Sierra's agents handle customer interactions end to end across chat, email, voice, SMS, and WhatsApp.

Sierra raised \$350 million at a \$10 billion valuation in 2025.

We include Sierra because it sits at the far end of the agentic pricing spectrum: fully autonomous, cross-functional, with an output-to-cost ratio that keeps widening. Sierra is the only company in our set that has fully committed to outcome-based pricing. They charge per successful resolution.

---

Between them, these five represent over \$60 billion in combined valuation and five pricing structures spanning from traditional per-seat to fully outcome-based.

## PART I

# The 5-Step Pricing Transformation Framework

Pricing is not about slapping a price point on a product. It is a strategic exercise that touches every part of your business.

At Monetizely, we use a 5-step framework that, when applied in order, takes you through the entire pricing strategy process:

### **Step 1: Goals & Segmentation**

Clarify what your company aims to achieve and identify who your customer segments are. Without alignment on strategy and ICP, pricing falls apart before it starts. Most pricing problems are actually alignment problems at this stage.

### **Step 2: Positioning & Packaging**

Design offers that fit your customer segments. Packaging is not bundling features into tiers. It is creating the right combination of capabilities for each segment so you capture the most value from each.

### **Step 3: Pricing Metric**

Select the variable that drives your pricing model. Per seat? Per usage? Per outcome? This is the most consequential decision in the entire exercise. Get it right, and everything else can be tuned. Get it wrong, and you slow down sales velocity and add friction that did not need to exist.

### **Step 4: Rate Setting**

Set the actual price point. This is a function of prospect willingness to pay, competitive and category price points, cost of delivery, and your strategic intent, whether you are optimizing for margin, revenue, or market share.

**Step 5: Operationalization**

The systems, processes, and people required to make the pricing model work day to day. The best pricing model in the world yields nothing if it is not built into a smooth sales process.

While all five steps matter, the three decisions that most directly determine whether your pricing model works or fails are Packaging (Step 2), Pricing Metric (Step 3), and Price Point (Step 4). Segmentation is the prerequisite. Operationalization is the execution. But the core architecture of your pricing lives in those three decisions.

For agentic AI, these three are where the most interesting tensions show up. Sequoia Capital's March 2026 piece "Services: The New Software" draws the line: a copilot sells the tool, an autopilot sells the work. That distinction reshapes every one of these decisions. A copilot packages like SaaS: tiers, seats, features. An autopilot packages like a service: outcomes, deliverables, capacity. Most of the five companies in this book sit somewhere in between, which is what makes their pricing decisions so worth studying.

## PART II

# Packaging, or Designing Offers That Map to Your Buyer

## The Fundamentals

Packaging is not about your product's features. It is not about your agent's autonomy level, its scope, or how much compute it burns. Those things matter later, when you pick a pricing metric and set a price point.

Packaging is about your customer segments.

The job of packaging is to create offers (the right combination of capabilities, services, and commercial terms) that map to the buyer segments in your market. Each segment has different needs, different willingness to pay, different buying processes, and different ideas about what "value" means. Your packages exist to capture the most possible value from within each segment.

The packaging approach lies on a spectrum:

Single Tier	Good Better Best (GBB)	Modular
<p>One plan, one price for all customers. Best suited for truly massive and uniform markets. For AI agents, this applies to internet-scale consumer products (personal AI assistants or tutors).</p>	<p>Graded tiers targeting different segments. The workhorse of SaaS and agentic AI packaging. Tiers differentiate by company size, buying complexity, or depth of need.</p>	<p>Bespoke, configurable approach. Suitable for mixed markets with varying industry needs and deal sizes. Requires consultative sales but captures higher revenue per account.</p>
<i>Massive uniform market</i>	<i>Large segmented market</i>	<i>Mixed specialized market</i>

The general rule: the more massive and uniform your market, the further you lean toward single tier. Larger but segmented, lean toward GBB. Smaller and more mixed, toward modular. Many companies blend approaches.

The mistake AI agent companies make more often than most is designing packaging around their product instead of around their buyer. They look at what the agent can do and build tiers around capability levels. But if those tiers do not map to real buyer segments with different needs and willingness to pay, you end up with packaging that confuses the market or leaves money on the table.

So for each of our five companies, we start with one question: Who are the customer segments, what are their jobs to be done, and does the packaging serve each one?

---

### **Cursor: Segments Understood, Capabilities Mapped Well**

Segment 1: The individual developer. Their job to be done is writing better code faster. They want AI autocomplete, multi-file editing, and codebase-aware suggestions. They do not need team management or admin dashboards.

Segment 2: The professional development team. Same core coding capability, but wrapped in a team environment. Centralized billing, admin controls, usage visibility. The AI features are the same. The organizational layer around them is what changes.

Segment 3: Enterprise engineering organizations. Adds a governance layer. SSO and audit trails, dedicated support, compliance certifications.

Cursor's packaging: Free / Pro / Business / Enterprise.

The mapping is clean. What Cursor got right is that the feature gates between tiers are not about rationing the AI. The core capability is largely the same across tiers. The tiers differentiate on the organizational and administrative features that each segment needs. This works because Cursor's market is large and the core job to be done is the same across buyers. That is the condition where GBB shines.

---

### **Devin: Right Segments, But the Packages Do Not Match Their Jobs**

Segment 1: The curious developer or small team. Their job is to test whether autonomous coding works for their kind of tasks.

Segment 2: The scaling engineering team. Their job is to add engineering capacity without adding headcount.

---

Segment 3: Enterprise engineering. Their job is deploying Devin as a capacity multiplier across many teams.

Devin's packaging: Core / Team / Enterprise.

The segments are right. But the packaging has a gap. Core includes about 9 ACUs, roughly enough for a handful of tasks. That is not enough for Segment 1's actual job to be done, which is running a real evaluation. Between Core and Team sits a real group: small teams of 3 to 10 engineers whose job to be done is running Devin on a few tickets a week. There is no package for them.

Cognition may have left this gap on purpose, funneling serious users to Team. But if that in-between segment is not served, competitors whose packaging does meet that need will capture those users.

---

## Harvey AI and Sierra AI: Right Capabilities for One Segment, Invisible to the Rest

Harvey and Sierra have made the same structural choice: build packaging that serves one premium segment and offer nothing to the others.

Harvey's Am Law 100 segment is well served by opaque, enterprise-only, bespoke deployment. But mid-size law firms (50 to 200 attorneys) and Fortune 500 in-house legal departments have different jobs to be done that require different capability bundles, and there is no package for either.

Sierra's large enterprise segment is well served by outcome-based, bespoke implementation. But mid-market companies wanting AI for routine support on one channel, and SMBs wanting a bot for their 20 most common questions, cannot engage with a package designed for multi-channel enterprise operations.

### COMPETITIVE RISK

For both Harvey and Sierra, as their markets get crowded, those underserved segments with unmet jobs to be done become the beachhead where competitors establish themselves. A mid-market firm whose job Harvey does not package for will buy from whoever does. And once a competitor owns that relationship, moving upmarket gets much easier.

## 11x (Alice): One Package That Speaks to No Segment's Job

Segment 1: Early-stage startups building outbound for the first time. Segment 2: Growth-stage companies scaling outbound. Segment 3: Enterprise sales organizations.

11x's packaging: one product, one offering, flat.

None of the three segments' jobs to be done are addressed by the packaging. Segment 1 needs a stripped-down, low-commitment entry. Segment 2 needs a package that positions Alice as a team member alongside human SDRs. Segment 3 needs enterprise controls and volume flexibility. In a market with 50+ competitors, failing to package for each segment's specific job means losing deals at the top, the bottom, and increasingly the middle.

### THE PACKAGING TAKEAWAY

Many conversations focus a whole lot on the pricing metric for an agent. But the basics of packaging have never changed and do not change even with Agentic AI products. This is where a lot of the money is made. Don't miss it.

## PART III

# Pricing Metric and Price Point

## Enter the Agentic Monetization Spectrum (AMS)

### Why We Built the AMS

Selecting your pricing metric is the most consequential decision in the entire pricing exercise. The right metric aligns your revenue with your customer's sense of value. The wrong metric either leaves money on the table or kills adoption.

For traditional SaaS, the evaluation criteria are well understood: Does the metric tie to client value? Do clients see it as fair? Is it measurable? Predictable? Do costs scale or shrink? Does it create workable deal economics?

These criteria still apply for AI agents. But agents have commercial properties (varying autonomy, scope, and output-to-cost ratios) that create questions traditional SaaS frameworks were not built for. That is why we developed the Agentic Monetization Spectrum (AMS). It does not replace the 5-step framework. It sits within Steps 3 and 4 as a classification system that maps your agent's properties to the right pricing metric and rate-setting approach.

### The Three Dimensions

#### **Dimension 1: Zero Human Ability**

*How much human involvement does the agent still need?*

Small (50%+ human, the human does the work, agent assists), Medium (20 to 50% human, human delegates, agent executes, human reviews), or Large (under 20% human, the agent does the work).

Why it matters: This is the strongest predictor of how the pricing model will be anchored. At Small, per-seat pricing still works because the human is the anchor. At Large, pricing must track the output or outcome the agent produces.

**Dimension 2: Operational Domain**

*How broad is the agent's scope?*

Small (single task), Medium (end-to-end workflow within one business function), or Large (cross-functional or multi-domain).

Why it matters: A narrow agent is a tool. A moderate agent is a mini job function. A broad agent is a department. The pricing metric has to match the buyer's mental model of what they are buying.

**Dimension 3: Output/Cost Curve**

*What is the ratio of output value to compute cost?*

Linear (output and cost move in lockstep), Inflecting (output begins to outpace cost, 10:1 to 100:1), or Exponential (output far outpaces cost, 100:1 to 10,000:1).

Why it matters: The curve tells the vendor how aggressively they can price. Linear means cost-based pricing works. Exponential means cost-based pricing makes no sense: it would price a \$5,000 deliverable at \$2.

**The Five Agents on the AMS**

Company	Zero Human	Domain	Output/Cost
Cursor	M	M	Inflecting
Devin	L	M	Inflecting to Exponential
Harvey AI	M	L	Exponential
11x (Alice)	L	M	Inflecting
Sierra AI	L	L	Exponential

**How the AMS Determines Pricing Metric**

Cursor (M / M / Inflecting)

Human is still the quality gate, single-function domain. Per seat is a natural fit. The buyer is purchasing productivity gains per developer.

Current metric: Per seat

Our assessment: Correct for its AMS position. The risk is that Cursor's agent capabilities outgrow per seat before the market is ready for something more complex. But getting the structure right at the root matters more than the price point.

### Devin (L / M / Inflecting to Exponential)

The agent does the work, but still within a single function. This is where per seat starts to feel like a ceiling and per output starts to feel possible. Cognition chose the usage path. The ACU model is a hybrid: part platform fee, part usage. The choice to anchor ACUs to compute cost rather than output value is on purpose. Devin's 15 to 30% success rate on complex tasks means value-anchored pricing would invite ROI scrutiny the reliability cannot yet survive.

Current metric: Hybrid: platform fee + usage (ACUs)

Our assessment: One of the best-designed pricing metrics in agentic AI. The base captures "access to an autonomous worker" value, while ACU consumption scales with actual work done.

### Harvey AI (M / L / Exponential)

The output dwarfs the cost. The Output/Cost Curve points clearly to value-based pricing. And yet Harvey uses per seat. Harvey's buyers are law firms whose entire commercial model is built around billable hours and seat-based software. They budget by headcount. Outcome pricing might be better economically, but it creates procurement friction in an industry that does not buy that way.

Current metric: Per seat

Our assessment: Not ideal for the AMS position, but defensible because of how law firms actually buy. Where Harvey should evolve is adding a usage layer on top.

### 11x / Alice (L / M / Inflecting)

The agent runs on its own, but output quality is uneven enough that the effective value ratio is not reliably high. The flat monthly fee creates misaligned incentives. If Alice underperforms, the buyer feels overcharged. If Alice overperforms, 11x captures the same revenue regardless. This is the gym membership model.

Current metric: Flat monthly fee

Our assessment: A hybrid model (lower base with a per-qualified-meeting kicker) would align incentives and give 11x a way to earn more from high-performing deployments.

### Sierra AI (L / L / Exponential)

Fully autonomous, cross-functional, with an output-to-cost ratio that keeps widening. A resolved customer interaction costs pennies in compute. The value of that resolution far outpaces cost.

Current metric: Per outcome (per successful resolution)

Our assessment: Correct for its AMS position. Outcome-based pricing is not a universal goal all agents should chase. It is a model that works when specific conditions are met: clear attribution, immediate measurement, and high per-outcome value. Sierra meets all three.

## How the AMS Informs Price Point

With the pricing metric selected, the AMS also guides rate setting. The pricing problem every agent company faces: the value gap between what an AI agent produces and what it costs to run is enormous and growing. But price is bounded by the lower of two ceilings: the value of the output, or the price of the next-best replacement. And that second ceiling is collapsing.

Company	Price Point	Strategic Intent
Cursor	\$20/mo Pro, \$40/user/mo Business	Market share - trade short-term revenue for workflow lock-in
Devin	\$20/mo entry, ACUs at \$2-\$2.25	Revenue optimization with path to margin as reliability improves
Harvey AI	~\$1,200/user/month	Margin optimization within a premium segment
11x (Alice)	~\$5,000/month	Margin, but under pressure
Sierra AI	Per resolution, ~\$150K-\$350K+ Year 1	Value capture at the enterprise level

## PART IV

# What We Have Learned

We applied the 5-step framework to five agent companies. Here is what holds up.

### **Packaging is still where the money is made.**

Cursor got it right by building tiers around buyer segments, not AI capability. Harvey and Sierra got it right for one segment and left the rest open for competitors. 11x got it wrong by offering one flat package that matches nobody's job to be done. None of these are technology problems. They are segmentation problems. Packaging has not changed because agents arrived.

### **The AMS maps agents to the right pricing metric.**

When the human does the work and the agent assists, per seat holds (Cursor). When the agent does the work in a single domain, a hybrid of platform fee plus usage works (Devin). When outcomes are immediate, attributable, and measurable, outcome pricing becomes possible (Sierra). But "possible" is not "advisable." Harvey sits at Exponential on the cost curve and still uses per seat because its buyers budget by headcount. The right metric is not the one that captures the most value in theory. It is the one your buyer will actually pay on.

### **Price is bounded by replacement cost, and replacement cost is collapsing.**

Inference drops 10x per year. Open source closes gaps. Every category floods with VC-backed competitors. Even agents at Exponential on the Output/Cost Curve may not capture much of the value gap. Cursor prices for market share. Devin prices on compute with a path to value. Harvey prices for margin. 11x is getting squeezed. Sierra bets that implementation depth is the moat, not the AI. Each is a bet on strategic intent. The price point is the last decision you make, not the first.

## What Comes Next

Most agent companies will land on hybrid pricing (platform fee plus usage or outcome component). Pure outcome pricing will stay rare because the preconditions are rare. The survivors in crowded categories will be the ones whose packaging evolved with their segments.

And services revenue will become core to how agent companies make money. We see it already with our clients. For a limited price of software, the real ticket is in the services wrapped around deploying it. That changes how you package, what metric you choose, and what price you can charge.

---

## The Agentic AI economy is here.

---



SaaS and AI Pricing strategy that consistently increases revenue by 12-40%.

Delivered via tech-enabled consulting in record speed.

[getmonetizely.com](https://getmonetizely.com)