

번역자료

(66)

310. 123  
9141

# 보스킨 위원회 보고서

## (The Boskin Commission Report)

2000 . 2



124289

6 3 4290

기 획 과

## 알 림

1. 보스킨 위원회의 공식명칭은 『소비자물가지수 연구를 위한 자문위원회』이며, 동위원회는 미국 상원 재정위원회에서 미국의 저명한 경제학자, 통계학자, 통계행정가들을 위원으로 임명하여 미국의 소비자물가지수 개선방안을 마련하기 위해 설립된 것입니다.
2. 1996년 12월에 발표된 보스킨 위원회 보고서는 미국의 소비자물가지수가 상향 바이어스되어 있음을 지적하여 큰 반향을 불러일으켰으며 미국 BLS에서는 이 보고서를 토대로 소비자물가지수 개선작업에 착수했습니다. 이 후 보스킨 위원회 보고서는 세계 각국에서 소비자 물가지수 개선 관련 업무에 참고자료로 널리 이용되고 있습니다.
3. 기획과에서는 소비자물가지수 품질평가를 위해 이 자료를 참고하였으며 우리청 지수 관련 업무 담당자의 편의를 위해 번역자료를 발간기로 하였습니다. 이 책의 번역을 담당한 기획과 정주용 사무관과 김신호씨의 노고에 감사드리며, 이 책이 널리 이용되길 기대합니다.

2000. 2. 11 .

기획과장 이 동 명

# 목 차

최종 요약 .....	3
I. 도 입 .....	9
II. 연방 예산의 연동 .....	14
III. CPI가 구성되는 방법 : 간략한 소개 .....	19
IV. 소비자물가지수와 생계비 지수 : 측정 이슈 .....	26
V. 품질 변화와 신상품 .....	38
VI. 각 종류별 및 총 바이어스 추정치 .....	68
VII. 기타 이슈 .....	75
VIII. 위원회의 권고안 .....	81
IX. 결 론 .....	91
원문 : The Boskin Commission Report .....	95

# 생계비의 보다 정확한 측정을 향하여

## 최종 요약

1. 미국 경제는 유연하고 역동적이다. 끊임없이 새로운 상품들이 등장하며, 기존 상품들이 향상되는 반면, 다른 상품들은 시장에서 사라진다. 서로 다른 상품 및 서비스의 상대가격은 소득과 비용 및 품질에 영향을 미치는 기술적 및 기타 요인들에 반응하여 종종 변동한다. 이는 정적인 경제에 비해 정확한 생계비 지표(cost of living index)의 구성을 곤란하도록 한다.

2. 생계비 지표를 추정하는 일에는 가정, 방법론, 자료수집 및 지수 숫자 구성 등이 필요하다. CPI의 강점은 그 개념에 내재된 단순성에 있다 : 고정된(그러나 대표적인) 상품 및 서비스의 장바구니(market basket)의 가격을 여러 시기에 걸쳐 조사하는 것이다. 그 약점은 동일한 개념에서 도출된다 : “고정된 장바구니”는 시간이 지남에 따라 소비자들이 가격변화와 새로운 선택가능성에 대해 대응하면서 점차 대표성을 잃게 된다.

3. CPI 변동을 생계비 변동의 척도로 사용하는 데에는 여러 개의 범주 또는 종류의 잠재적 편의가 존재한다. 1) 대체 바이어스(substitution bias)가 발생하는데 이는 상대가격이 변화할 경우 보다 값비싼 상품으로 교체하는 소비자의 비율이 상대적으로 줄어든다는 사실을, 고정된 장바구니는 반영하지 못하기 때문이다. 2) 상점 바이어스(outlet bias)는 가격이 보다 저렴한 할인점으로 전환하는 현상을 적절히 처리하지 못할 경우 발생한다. 3) 품질 바이어스는 상품 품질의 개선, 예를 들어

에너지 효율성 향상 또는 수리 필요성의 감소 등이 부정확하게 측정되었거나 전혀 측정되지 않을 경우 발생한다. 4) 신상품 바이어스는 새로운 상품들이 장바구니에 도입되지 않거나, 오랜 지체 끝에 비로소 도입될 경우 발생한다.

4. CPI가 현재 이용 가능한 것 중에서는 최고의 척도임에도 불구하고, 진정한 생계비 지표는 아니다(이는 노동통계국이 오래 전부터 인정해 온 사실이다). 노동통계국이 CPI에 여러 개의 중요한 업데이트 및 개선을 가했음에도 불구하고, CPI 변동폭은 앞으로 몇 년간 진정한 생계비 변화를 과대 평가할 것이다. 향후 상향 바이어스(upward bias)에 대한 위원회의 점추정치는 1년에 1.1 퍼센트 포인트이다. 구간 추정치는 1년에 0.8에서 1.6퍼센트 포인트 사이이다.

5. CPI변동폭은 실제 물가 인플레이션율을, 1996년 이전 1년에 1.3% 포인트로, 상당히 과대 평가해왔다(추가 0.2% 포인트는 산식 바이어스(formula bias)로 지칭되는 것에 기인하는데, 이는 1978년에 우연히 도입되었으며 올해 들어 정정됨). 적어도 지난 20년을 되돌아볼 때에도 대규모의 바이어스가 일어났을 가능성이 높다.

6. 상향 바이어스는 연간 물가연동 이익 증가와 실질조세 감소를 연방 예산에 가져온다. CBO는 만약 앞으로 10년에 걸쳐 CPI변동이 생계비를 평균 1년에 1.1 퍼센트 포인트 과대 평가할 경우, 이 바이어스는 2006년 적자에 대략 \$148 billion를, 국가부채에 \$691 billion를 더할 것으로 추정하고 있다. 이 바이어스 자체만으로도 사회보장, 보건 및 국방에 이어 규모순위가 4위인 연방 프로그램이 될 것이다. 2008년에 이르면, 이 합계들은 차례대로 \$202 billion와 \$1.07 trillion에 이를 것이다.

7. 일부에서는 인구내의 서로 다른 집단들이 CPI변동으로 기록된 것보다 빠른 또는 느린 속도의 생계비증가를 경험할 가능성이 높은 것으로 주장하였다. 현재에 이르기까지 이를 위한 강력한 증거를 발견하지는 못했으나(실제는 오히려 정반대였음), 이 이슈를 심층 탐구하는 것은 바람직하다.

8. 위원회는 노동통계국에게 대략 12개의 권고사항들을 넘겨줄 것이다. 이는 다음을 포함한다.

i. 노동통계국은 소비자물가 측정에 있어서 생계비 지표(Cost of Living Index, COLI)를 그 목표로 확립해야 한다.

ii. 노동통계국은 2개의 지표를 개발 및 발표해야 한다 : 한 개는 매월 공표하며, 한 개는 매년 공표 및 업데이트되고 소급하여 개정한다.

iii. 시의성 있는, 월별 지수는 계속 CPI로 지칭하고, “최고(superlative)” 지표 산식을 채택함으로써 변동하는 장바구니를 설명하기 위해 COLI 개념을 지향하면서, 고정가중치 라스파이레스 산식을 유지한다는 구실을 타파해야 한다.

iv. 새로운 연간 생계비 지표는 호환성 있는 “최고 지표” 산식을 이용하고, 이에 따른 자료, 업데이트된 가중치, 그리고 신규 상품의 등장(그 역사를 과거로 연장한 채)을 반영할 것이다.

v. 노동통계국은 기초합계(elementary aggregates) 수준에서 기하평균(geometric means)으로 전환함으로써 가격 관측치 통합절차를 변경해야 한다.

vi. 노동통계국은 개별 지표 구성요소들의 행태를 관찰하여 지표의 향후 장기추세에 대해 가장 정보를 많이 주는 요소들이 어떤 것인지, 그리고 총합과 대체로 무관한 동요를 보이는 요소는 어떤 것인지를 확정짓고 전자를 자료수집 활동에서 강조해야 할 것이다.

vii. 노동통계국은 지역에 대한 고려를 줄이는 방향으로 소비자물가지수 표본추출절차를 변경해야 한다. 우선, 가격이 조사되는 상품 모집단을 추출한 후, 상품별로 어떠한 방식으로, 어떠한 상점에서 수행하는 것이 대표적인 가격표본을 수집하기 위해 가장 효율적인지 결정해야 한다. 그 후에 비로소 자료수집의 경제성을 위해 지리적으로 군집된 표본에 의존해야 한다.

viii. 노동통계국은 품목군의 정의 및 구조를 뜻하는 분류가 물가지수에 미치는 파장(impact)을 조사하여 지수가 품목대체를 전적으로 반영할 수 있도록 개선해야 한다.

ix. 주의를 요하는 개념적 이슈가 몇 가지 더 있다. 자동차와 같은 내구재 가격은 자가소유주택의 가격에 대한 현행 처리방법과 동일선상에서, 연간 서비스의 가격으로 변경해야 한다. 또한, “보험”의 처리도 현재 사용되는 보험수익에 기초한 사후적 측정수단(ex-post insurance profits based measure)보다는 사전적인 소비자물가 측정수단(ex-ante consumer price measure)으로 이행해야 한다.

x. 노동통계국은 외부의 정보, 전문지식 및 연구결과를 도입하기 위한 영구적 메커니즘이 필요하다. 노동통계국 요청에 의하여, 이 단체는 독립된 공공의 전문가집단(professional entity)으로 조직되어야 하며, 노동통계국이 통계, 경제 및 최신 시장이슈에 관한 전문가 및 사업체 소견에 접하는 통로를 개선할 것이다.

xi. 노동통계국은 소비자물가지수의 현행 “장바구니” 구조를 극복할 수 있도록 연구프로그램을 개발해야 한다.

xii. 노동통계국은 균형가격의 가정이 깊이 각인된 데에 따르는 과장과 0이것이 때로는 성립하지 않는 경우에 대한 함의를 조사해야 한다.

xiii. 노동통계국은 이와 동일선상에서 진전을 이루기 위해 몇 가지 신규 자료수집 이니셔티브를 취할 필요가 있다. 가장 중요한 것으로, 대규모 소비자표본으로부터 세부 시간사용에 대한 자료가 개발되어야 한다.

9. 위원회는 대통령 및 의회에 여러 개의 권고를 하고 있다. 이는 다음을포함한다 :

xiv. 의회는 경제통계의 정확성과 시의성을 증진시킬 목적으로 정보를 공유하기 위해, 그리고 개발 및 생산과정에서 소요되는 자원을 절감하기 위하여, 상무부와 노동부가 필요로 하는 법령을 제정해야 한다.

xv. 의회는 CES 표본과 수집되는 자료항목을 확대하고, POPS 조사의 수행빈도를 높이고, 산업조사 및 스캐너자료와 같은 여타 국내 원천으로부터 상품 상세정보를 더 많이 획득하는 데에 필요한 자원을 추가로 제공해야 한다.

xvi. 의회는 이 분야의 발전을 3년여 주기로 점검하고 통계현황에 대해 적절한 해석을 자문하기 위하여 상설(순환) 독립 위원회 또는 전문 위원단을 설립해야 한다.

xvii. 의회와 대통령은 다양한 연방지출 프로그램과 조세법령 규정에 널리 퍼져있는, 현저한 과대 물가연동(overindexing)이 지속되는 것을 바라는지 결정해야 한다. 만약 물가연동의 목적이 이전지출 수렴집단과 납세집단을 정확하고 완전하게 분리하는 데에 있으며, 그 이상도 그 이하도 아니라면, 의회는 물가연동 규정을 이에 맞게 조정한 법률을 통과시켜야 한다.

이는 현행 소비자물가지수에 연동한 값으로부터, 과다연동을 부분적으로 또는 전적으로 반영한 액수를 공제한다는 맥락에서 행할 수 있다. 다른 대안으로는, 만약 새로 개편된 연간지수를 개발하고 정기적으로 공표할 경우, 동 지수에 대해서는 보다 작은 액수를 공제하여 생계비 변동에 보다 근접하도록 할 필요가 있을 것이다.

우리는 물가연동 프로그램에, 보다 정확한 생계비지표라는 협소한 범위를 초월한 여러 개의 다른 측면들이 있으며, 여러 개의 다른 이슈들이 제기된다는 사실을 서둘러 덧붙이고자 한다. 우리는 또한 노동통계국, 의회, 행정부(executive branch), 그리고 국민들이 이 견해 및 그 함의를 완전히 숙고하고 또 이해할 필요가 있다는 사실을 표명하고자 한다.

## 1. 도입

생계비변동을 정확히 측정할 수 있는 수단들은 경제성과를 평가하는데에 필요한 가장 유용하고도 중요한 정보에 속한다. 예를 들어 1975년과 1995년 사이와 같은, 2개의 시기간 생계비 변동은, 1995년 당시 소득과 이용 가능한 상품 및 서비스의 가격이 주어졌다고 가정하고, 1975년의 상품 및 서비스의 가격이 주어졌다고 가정할 때, 1995년과 적어도 동등한 생활수준을 누리기 위해 1975년에는 얼마나 많은 소득이 필요했는지 우리에게 알려준다. 예를 들어, 1996년 소득이 \$45,000인 가족이 동일한 후생수준을 누리기 위해 1976년에는 \$15,000가 필요했다면, 그 사이에 생계비가 세배수로 증가한 것이다.

미국 경제가 만약 신상품이 거의 등장하지 않고, 기존 상품의 품질향상이 매우 드물고, 소비자 소득변화가 적고, 또한 상품 및 서비스의 상대가격 변동이 빈도와 규모 면에서 모두 작은, 상당히 정적인 경제였다면, 생계비 변동의 측정이 개념적으로 상당히 쉬웠을 것이며 이를 이행하는 것 또한 기술적인 세부사항과 적절한 실행의 문제였을 것이다. 미국인들의 압도적인 다수를 위해서는 다행히도, 우리의 경제가 이보다 월등히 역동적이고 유연성 있다. 신상품이 끊임없이 등장하고, 기존 상품들은 개선되는 반면, 다른 상품들은 시장에서 사라진다. 서로 다른 상품 및 서비스의 상대가격은 소비자 수요, 그리고 비용 및 품질을 좌우하는 기술 및 기타 요인의 변화에 대응하여 빈번히 변동한다. 미국의 소비자들은, 일부 여타 국가들 또는 수십 년 전 우리의 조상들에 비해, 광대하고 계속 증가하고 있는 상품 및 서비스 배열로부터 선택할 수 있는 혜택을 누리고 있다.

그러나 복잡하고 역동적인 경제로 인해 정확한 생계비지표를 개발하는 데에 보다 큰 어려움이 있는 것을 개탄할 이유는 못된다. 유명한 소비자물가지수(CPI)와 같이, 현재 생계비변동의 대리변수(proxy)로 사용되고 있는 다양한 공식통계는 크게 개선될 수 있으며, 또 개선되어야 한다.

소비자물가지수는 고정된 상품 및 서비스의 장바구니 구입에 드는 비용을 측정한다. 어떤 기준시기의 가구조사를 바탕으로 하여 이 지수는 서로 다른 상품 및 서비스의 가중치(지출비중)를 설정한다. 가중치는 조사집단의 평균 또는 대표적 비중을 나타낸다. 이 가중치를 계속 고정시킨 채, 한 달로부터 다음 달에 이르기까지 동일한, 또는 밀접히 관련된 상품 및 서비스의 가격변동을 측정하려고 시도함으로써 CPI가 산출된다.

그러나 시간이 지나면서 장바구니는 변화하는데, 이는 부분적으로 상품 및 서비스의 상대가격 변화에 기인한다. 따라서 기준시점의 가중치는 더 이상 소비자들이 실제로 구입하는 것을 반영하지 못한다. 대표성 있는 구매는 또한 할인쿠폰, 회원제 클럽(buyer's clubs) 그리고 기타 마케팅 기법들이 최고 가치를 결정하고 구매패턴을 변형시킴에 따라 변화한다. 상대가격 변화에 대응한 소비자의 행태변화에 대해 지수가 적응하지 못하는 이 현상을 대체바이어스(substitution bias)라고 한다. 이는 장바구니를 계속 고정으로 유지하는 데에 따른 불가피한 결과이다. 장바구니는 10년에 한 번 정도 업데이트되므로, 기준시점에서 멀리 떨어질수록 상대가격이 기준시점의 값을 벗어나 있으며, 장바구니가 현저히 변화했을 가능성이 높다.

소비자들이 구입하는 물건이 변화하는 것과 마찬가지로, 구매가 이루어지는 장소에 또한 추세와 변동이 존재한다. 최근 몇 년 동안, 소매업에서 변환이 일어났다. 대형 백화점, 할인점 및 기타 등등 대략 20년 전에 비해 매출 중 매우 높은 비중을 차지하고 있으며, 이 비중은 점차 증가하고 있다. 소비자들이 실제 구입하는 상품바구니를 놓치지 않고 따라가는 것만큼이나 중요한 것은 소비자들이 실제로 이 상품바구니를 구입하는 상점들을 놓치지 않고 따라감으로써, 지불된 가격을 정확하게 기록하는 일이다. 현행 방법론은 할인 상점으로의 전환을 충분히 반영하지 못하는, 상점 대체 바이어스 하에 있다.

오늘날 판매되는 상품중 상당수는 수년 전의 판매되던 해당 상품의 형태(counterparts)를 엄청나게 개선한 것들이다. 이들은 보다 내구적이고, 수리가 덜 요구되며, 보다 에너지 효율적이거나, 가볍거나, 안전하거나, 등등일 수 있다. 때때로, 적어도 애초에는, 이전의 형태를 교체하는, 보다 우수한 품질의 상품에 보다 많은 비용이 들 수 있다. 가격 상승 중 어떤 규격화된 상품의 실제 인플레이션보다는 품질변동에 기인한 부분이 어느 정도인지 분리해내는 일은 단순함과 거리가 멀다. 그러나 이는 진정한 생계비 상승에 대한 정확한 측정치를 얻기 위해 필요하다. 품질변동이 부정확하게 측정되거나 아예 측정되지 않고 있는 한, CPI에 품질변동 바이어스가 존재한다.

이는 그 자체로 상당한 가치를 보유하고 있을 뿐 아니라, - 우리들 중 전자레인지, 레이디얼 타이어(radial tires), 그리고 VCR을 기꺼이 포기할 사람은 많지 않다 - 이로 인해 열리게 된, 보다 넓은 선택 및 기회 가치도 보유하고 있는 신상품이 소개될 때에도 마찬가지로 적용된다. 신상품이 장바구니에 포함되지 않거나, 오랜 지체 끝에 비로소 반영되는 한, CPI에는 신상품 바이어스가 존재한다.

마지막으로, 현대의 미국과 같은 역동적이고, 복잡한 경제 속에서는 말 그대로 수천 개의 상품 및 서비스가 소비된다. 물가자료는 상당히 세부적인(disaggregation) 수준에서 수집된다. 그리고 가격변화가 하나의 전체 지수로 합쳐지는 방법은 상당히 기술적인 이슈들을 수반하는데, 이는 CPI에 산식 바이어스를 가져올 수 있다.

심지어 예산지출과 수입측면에 속한 연방 프로그램이 모두 물가에 연동되지 않는 경우에도, 생계비척도의 품질을 개선하는 것은 다음의 측면에서 바람직하다. 즉, 이를 통해 시민들에게 경제에 실제로 어떤 일이 일어났는가에 대해 보다 나은, 그리고 보다 정확한 추정치, 현재 경제성결과 과거성과 또는 여타 국가들의 성과와 비교하기 위한 방안을 제공할 수 있기 때문이다. 예를 들어, 가장 흔히 사용되는 생활수준의 척도는 실질소득 또는 1인당 산출이다. 실질소득 변동을 측정하기 위해서는 명목소득을 물가 변동으로부터 분리해내는 일이 필요하다. 이를 위해 물가변동의 정확한 척도가 필요하다는 사실은 명백하다. 상무부는 인플레이션 및 실질GDP 추정의 투입요소로서 CPI 구성지표들을 사용하고 있으며, 따라서 CPI의 바이어스 중 몇 가지는 국민소득계정에 이전되고 있다.

그러나 수많은 연방, 주(州) 및 지역정부 프로그램과 조세항목(tax features)은 소비자물가지수 변동을 통해 생계비 변동에 연동(indexed)되어 있다. CPI는 또한 공식적으로 또는 비공식적으로, 상당수의 민간부문 계약들을 연동시키기 위해서도 이용된다. 수백만 명의 미국인들에게 영향을 미치는 명백한 예를 몇 가지 언급하자면, 집단적 교섭 협정(collective bargaining agreements)을 통한 임금의 결정, 임대료 등이 있다. 현재, 대부분 연금프로그램에 해당하는, 전체 연방지출 중 1/3을 약간 못 미치는 비율이 소비자물가 변화에 연동되어 있다. 소득구간을 포함한, 개인 소득세의 여러 가지 항목들 또한 연동되어 있다; 개인소득세는 연방 조세수입의 절반에 약간 못 미치는 비율을 차지하고 있다.

의회는 그 영향하에 놓인 개인들이 생계비 가중부담을 정면으로 받는 데에 대한 방지책 또는 보호책으로서, 이 지출프로그램과 조세규정을 연동시킨 바 있다. 그럼에도 불구하고 소비자물가지수의 축적 및 공표를 담당하고 있는 노동통계국은, 수년간 CPI가 생계비지표가 아님을 선언해왔다. 이는 위에서 언급된 이유 중 일부에 기인한 것으로 추측된다. 만약 현재 작성되고 있는 방식의, 그리고 앞으로 몇 년간 그렇게 작성될 가능성이 높은 소비자물가지수가, 의회에서 염두에 두고 있는 과업을 위한 적절한 생계비 지표가 아니라면, 여타 방안을 생각해 보는 것이 바람직하다.

소비자물가지수 변동이 생계비변화를 과대 평가하는 데에 따르는 결과는 엄청날 수 있다. 예를 들어, 앞으로 12년간 만약 CPI 사용이 생계비 증가를 1년에 1퍼센트 포인트씩 과대 평가할 것으로 예상된다면, 지출과 수입의 물가연동을 적절히 정정했을 경우에 비해 2008년 국가 부채가 대략 \$1 trillion 더 증가할 것이다.

이 보고서는 다음과 같이 전개된다 : 제Ⅱ부는 CPI변동이 생계비변동을 과대 평가하는 데에 따른 과거 및 미래의 예산상 함의를 논의한다. 제Ⅲ부는 CPI가 실제 어떻게 구성되는지에 대한 개관을 제시한다. 제Ⅳ부는 CPI가 진정한 생계비지수가 되지 못하는 이유를 세부적으로 설명하고, 대체 바이어스를 논의한다. 제Ⅴ부는 현재 품질변화에 대한 조정을 하기 위해 노동통계국이 현재 채택한 절차들을 보다 세부적으로 기술하고, 품질변화와 신상품으로부터 나온 바이어스에 대하여 연구조사 및 위원회의 판단을 설명한다. 제Ⅵ부는 유형별 바이어스의 규모, 그리고 전체 바이어스의 구간추정치에 대한 위원회의 연구 결과를 요약하고 있다. 제Ⅶ부는 서로 다른 집합과, 시장에 바탕을 둔 소비-관점의 생계비 척도에서 주로 벗어난 경우가 많은 ‘삶의 질’의 속성들을

위한 별도의 물가지수 관련 이슈를 논의하고 있다. 제Ⅷ부는 보다 정확한 생계비척도를 생산 및 사용하기 위한 방법들에 대한 위원회의 상세한 권고사항들을 제시한다. 결론에서는 간략한 전망과 위원회 연구결과 이용시 몇 가지 주의사항이 제공된다.

## II. 연방 예산의 연동

CPI의 상향 바이어스로 인해 재정정책 담당관들이 직면하게 된 이슈를 의회예산처(Congressional Budget Office, 1994)는 감탄할 정도로 명쾌하게 선언한 바 있다 :

생계비 변동이 과대 평가되는 경우 항상 예산으로 나타나는 그 효과로 인해 부(富)의 분포가 이동할 가능성이 강조된다. 만약 CPI가 상향 바이어스를 보인다면, 일부 연방프로그램들은 물가변동이 생활수준에 미치는 영향을 과잉 보상하게 되고, 부가 젊은 미래 세대로부터 물가 연동 연방 프로그램의 현재 수혜자들에게 이전될 수 있다. 그리고 이는 입법의도와 다를 수 있다.

사회보장이 CPI에 연동된 연방지출 중 단연코 가장 중요하다. 그러나, 추가적 소득보조(Supplemental Security Income), 군 제대, 그리고 공무원 퇴직 등도 유사한 방식으로 연동되는 대표적인 프로그램들이다. 기타 연방 퇴직프로그램들, 철도 퇴직, 퇴역군인 보상 및 연금, 그리고 연방 고용보상법(Federal Employees' Compensation Act) 또한 연동을 위한 규정을 포함하고 있다. 1981년 경제회복법(Economic Recovery Act)은 개인 소득세 구간과 개인공제를 CPI에 연동시켰다.

역사적으로 CPI 상향바이어스의 예산상 결과는 얼마나 중요했는가? 이 질문에 대한 정확한 답을 얻기 위해서는 바이어스의 시점, 특정 연방지출 및 수입에 대한 연동규정의 평행선상 발전, 그리고 이로 인해 초래된 부채 축적의 이자 등을 모두 고려한, 광범위한 연구가 필요할 것임이 명백하다. 이 효과의 잠재적 규모에 관한 암시는 명백히 확인된 하나의 바이어스 원천에 대한 하나의 중요한 역사적 실례로부터 유추할 수 있다. 물가연동의 영향을 받는 가장 중요한 연방 프로그램, 즉 사회보장 급여에 초점을 맞춘, 동종의 세밀한 연구가 재무부 경제정책청(Office of Economic Policy, OEP)에 의해 수행되었다.

1983년 2월 25일, 노동통계국은 모든 도시소비자 소비자물가지수(Consumer Price Index for All Urban Consumers, CPI-U)에 대한 중요한 기술적 정정을 도입했다. 이로써 자가소유자들의 주거비용 처리가 임대료-등가 바탕으로 변경되었다. 주거비용에 대한 새로운 처리방식은 도시 임금노동자 및 사무직 종사자(Urban Wage Earners and Clerical Workers) 소비자물가지수에 통합되어, 1985년, 사회보장 급여에 물가를 연동시키기 위해 이용되었다.

주거비용의 임대료-등가 척도는 개념상 진보였으며, 뒤이은 공식 간행물들에 계속 유지되어왔다. 그러나, 이전 연도의 주거비용은 주택가격, 저당 이자율, 자산세와 보험, 그리고 유지비용 등에 바탕을 둔, “주거소유” 방식을 채택하고 있었다. 1983년 이전의 주거비용 처리는 개정된 CPI-U를 공표하는 데에 있어서 수정되지 않았다. 이에 따라 1983년 전후 CPI-U의 개념적 바탕에는 불일치가 발생하게 되었다. 이와 유사한 방식으로, 1985년 이전 CPI-W의 주거비용 또한 수정되지 않았다.

노동통계국은 1967년에 이르기까지 소급하여 주거비용을 임대료-등가로 취급한 “실험적” 물가지수, CPI-U X1을 개발했다. 이는 경제정책청이 CPI-W의 바이어스를 평가하기 위한 기반을 제공하고 있다. 1975년 사회보장이 CPI-W에 연동되던 첫 해의 바이어스는 1.1퍼센트였다. 이 바이어스는 이후 수년간 증가하여, 1982년에 6.5%에 달했으며, 그 후 1984년에는 4.7%로 하락하였다.

CPI-W의 바이어스가 초래한 사회보장급여의 과잉지출은 1983년에 이르기까지 증가하여, 총 \$8.76 billion 또는 그 해에 지출된 사회보장급여의 5.55%에 이르렀다. 이 과잉지출은 OASI 신탁자금의 수지균형(balance) 하락과 연방 재정적자 및 부채증가를 초래했다. 경제정책청은 OASI 신탁자금에 대해 현재 지불되고 있거나 앞으로 지불이 예상되는 이자율 수준에서, 이 재정적자와 관련된 이자비용을 예측하고 있다. 1984년 이래로 이자비용은 전체에서 지배적인 부분을 차지하고 있다. 현 회계연도에 총 비용은 \$21.79 billion이며, 이 중 \$17.64 billion은 이자이다. CPI-W 바이어스의 이 원천 단 하나만으로, 이 프로그램 하나를 통해 연방부채에 미치는 누적적 효과는 1996년 현재, \$271.0 billion에 달한다.

요약하여, 노동통계국은 1985년에 CPI-W의 주거비용 취급방식을 개정하는 2가지 결정을 내렸다. 첫 번째 결정은 1985년 1월부터 주거비용 취급방식을 임대료 등가로 변경하는 것이었다. 둘째 결정은 1984년과 그 이전 기간에 대해 주거비용 취급방식을 개정하지 않는 것이었다. 이 2개의 결정 결과, CPI-W수준은 CPI-U X1에 비해 4.7 퍼센트 높은데, CPI-U X1은 CPI-W와 동일한 원시자료 및 유사한 방법론을 바탕으로 하였으나, 주거비용을 일관되게 취급한 생계비 척도이다.

CPI-W 바이어스로 인해 초래된 연방지출 증가는 생계비 조정으로 정당화될 수 없다. 이 증가는 1985년 이전에 주거비용이 부적절하게 취급된 결과이며, OASI프로그램 수혜자들에게 경제적 합당성을 전적으로 결여한 대규모 이전을 초래하였다. 과잉지출은 현재에 이르도록 계속되어 왔으나, 점차 그 중요성이 약화되고 있다. 그럼에도 불구하고, 이에 따른 OASI신탁자금 감소는 이자비용의 상승으로 인해 계속 쌓이고 있으며, 현재 연방부채의 \$200 billion 넘는 부분을 차지하고 있다.

물론, 누구도 이것 또는 다른 어떤 바이어스 원인에 기인한 과다연동을 소급하여 원상태로 되돌리도록 주장하지는 않을 것이다. 이 논의의 핵심은 그 결과가 쌓이고, 과연 심하게 가중되기 이전에, CPI의 바이어스를 가능한 한 빠르고 완벽하게 시정하는 것이 얼마나 중요한지 보여주는 것이다.

CPI의 상향 바이어스가 미래 예산적자에 미치는 영향은 어떠한가? 현재 \$1.5 trillion 연방지출의 절반 이상이 권리(entitlements)와 의무지출 프로그램에 기인한다. 1995년 1월, 경제와 연방예산에 대한 의회예산처 연간전망은 이 비율이 1998 회계연도에 연방지출의 거의 2/3 수준으로 상승할 것임을 보였다. 생계비를 가령 3.0% 비율로 시정할 경우 이로 인해 그 해의 의무적인 지출프로그램의 총지출이 \$43 billion 상승하게 되며, 회계연도 2000년도에는 \$80 billion 상승하게 된다. 이는 의무적인 프로그램에 대한 2000회계연도 예상지출의 6.8%이다.

의회예산처가 재정위원회(Committee on Finance)에 제출한 성명서는 회계연도 1996-2000에 대해 가상적으로 생계비 조정을 0.5퍼센트 포인트 시정(삭감)했을 때 나타나는 효과를 보여주고 있다. 2000 회계연도에 연방지출은 \$13.3 billion 감소하는 반면, 수입은 \$9.6 billion 증가할 것

이다. 회계연도 1996-2000동안 적자감소에 기인한 debt service의 감소는 \$3.3billion에 달할 것이며, 이는 회계연도 2000년도 적자감소에 총 \$26.2 billion을 기여하는 결과를 가져올 것이다. 이는 해당 연도의 의회예산처 예상적자의 10%를 초과한다.

의회예산처는 회계연도 1997-2006년도에 대한 가상적인 0.5 및 1.0% 포인트의 생계비 시정(감축)효과에 대한 예측치를 업데이트하여 위원회에게 제출하였다. 0.5% 포인트 감축을 통해 2006년 적자는 총 \$67.5 billion감소한다. 이 액수 중, 수입계정의 증가가 \$22.3 billion를 차지하고, debt service를 포함한, 지출의 감소가 \$45.3 billion에 달한다(이 중 debt service는 \$13.1 billion이다).

생계비 조정을 1.0% 포인트 시정(감소)했다고 가정할 경우, 그 효과에 대한 의회예산처 예측치는 물론 더욱 극단적이다. 2006년도 적자의 총 변동은 \$134.9 billion이다. 연방수입은 \$44.5 billion 증가하며, 연방지출은 \$90.5 billion 감소할 것이다; 지출감소분 중 \$26.1 billion은 debt service 감소에 기인하고, \$64.4 billion은 연동된 프로그램의 지출 감소에 기인한다(세부사항은 부록 그림 A-1 참조).

이를 달리 표현하여, 만약 이 기간에 걸쳐 CPI 변동이 생계비 변동을 1년에 평균 1% 포인트 과대평가 했다면, 이 바이어스만으로도 2006년도 적자가 거의 \$135 billion 증가할 것이다. 이는 기준적자(baseline deficit, 이는 현행 균형예산 제안 등의 정책변동을 가정하지 않는다)의 1/3에 해당한다. 보다 주의를 끄는 것은, 상향 바이어스가 그 자체로서 사회보장, 보건 및 국방에 이어 연방지출 프로그램 중 4위의 규모를 기록하게 된다는 것이다. 2008년에 이르면, 증가된 적자는 \$180 billion, 국가부채는 \$1 trillion에 달할 것이다.

요약하면, CPI의 상향 바이어스는 연방 권리(entitlements) 및 의무 지출 프로그램 수혜자들에 대한 상당한 과잉지출을 가져올 것이다. 그 외에도, 이러한 바이어스는 개인소득세를 과다 연동시킴으로써 연방수입 감소를 가져올 것이다. 한마디로, 상향 바이어스는 연방예산에 매년 물가연동 급여의자동적인 실질 증가와 실질 조세감면을 예정해놓고 있다. CPI 바이어스의 수정은, 급여 및 조세가 진정한 생계비 변동을 보다 정확하게 반영하도록 이를 조정하기 위해 설계된 동시에, 미래의 연방재정적자 및 국채 감소에 중대한 기여를 할 것이다. 이는 수입증가, 지출감소 및 debt service의 감소등에 그 원인을 돌릴 수 있다. 지출 하락 - 물가연동 연방지출 프로그램 감축 및 이자지출 감소 - 에 의해 장기 적자감소의 2/3을 넘는 비율이 설명되는 반면, 나머지 부분은 수입 증가에 의해 설명된다.

### III. CPI가 구성되는 방법 : 간략한 소개

바이어스가 발생하는 이유와, 개선 및 변경 권고의 근거를 이해하기 위해서는 CPI가 어떻게 구성되는지에 대한 지식이 필요하다. 이 부분은 바이어스와 주요 이슈들이 제기될 가능성이 높은 대목들을 강조하면서, 노동통계국 방법론에 대해 간략히 서술하고 있다. 자료수집절차와 지표구성 방법론뿐만 아니라, Armknecht(1996), Shapiro 및 Wilcox(1996b) 등의 최근 저술 등에 관한 세부사항은 노동통계국 발간 자료를 참조하도록 권장한다.

역동적인 현대 자본주의경제의 복잡성에 관한 위의 논의로부터 유추할 수 있는 바와 같이, CPI 프로그램은 복잡하고 어려운 사업이다. 그 운영이 가능하도록, 노동통계국은 시장(marketplace) 및 소비자 행동에

대한 단순화된 시각을 적용한다. 단순화된 이 시각은 CPI의 접근방법 전반에 걸쳐 나타난다. 이는 기준 또는 참조시점으로 지칭되는 과거 어떤 시점에 대해, 상품 및 서비스의 고정된 장바구니에 대한 지출을 산출한 후, 오늘날 동일한 장바구니를 구입하기 위해 드는 비용이 얼마인지 추정한다. CPI 구성에 사용되는 산식은, 라스파이레스로 지칭되며, 구매가 어떤 이전시기의 경험으로부터 나온 결정에 바탕을 둔 고정수량으로 이루어진다고 가정한다. 즉, CPI는 “이번 달의 시장가격으로, 기준시점에 실제 구입한 것과 동일한 장바구니를 구입하는 데 드는 비용은 얼마인가?”의 질문을 답하고자 한다. 라스파이레스산식이 현재가격과 선택에 대응하여 상품 및 서비스가 대체되는 것을 참작하지 않으므로, 이는 생계비의 “상한선”이다.

장바구니는 모든 도시 소비자들이 직접 구입하는 품목에 대한 총지출로 구성되어 있다. 즉 이는 식품, 의복, 주거 및 연료, 교통, 의료서비스와 하루하루 생계를 위해 사람들이 구입하는 기타 상품 및 서비스 등을 의미한다. 노동통계국은 특정 품목의 선정을 위해 과학적인 표본추출 기법들을 이용한다. 노동통계국은 이 품목들의 가격변화를 여러 시기에 걸쳐 측정한다. 표본설계는 지리적 지역, 소매상점, 품목 범주, 그리고 하나의 상점과 범주 내의 개별 상품 및 서비스에 대한 다단계 표본추출절차를 그 내용으로 한다.

CPI가 소비자 지불가격에 대한 대표성을 갖도록 하기 위해 여러 개의 표본들이 이용된다 : 미국내의 모든 도심지역으로부터 선정된 도심지역, 각 선택지역내의 소비주체(consumer unit), 이 소비주체가 상품 및 서비스를 구매한 상점들, 이 소비주체가 구매한 특정 품목 - 상품 및 서비스 - , 그리고 각 도심지역의 주거단위(housing units, CPI의 주거(shelter) 구성요소를 위함). 장바구니를 구성하는 품목과 가격이 수집

되는 상점을 결정하기 위해 사용되는 주요 정보원천은 소비지출조사(consumer Expenditure Survey, CES)와 구매시점조사(Point-of-Purchase Survey, POPS)이다.

매월, 대략 71,000개의 상품 및 서비스 가격이 44개 지역, 22,000개 상점에서 수집된다. CPI 주거요소를 위한 정보는 매월 대략 5,000명의 주택임대자와 1,000명의 자가소유자들로부터 별도로 수집된다. 가격측정치는 전체 CPI로 통합, 즉 집계된다. 가격이 조사될 대표품목의 결정, 상점에서 가격수집이 이루어지는 절차, 그리고 가격들이 지표로 통합되고, 지표들이 보다 상위 총계로 통합되는 수준은, 모두 몇 가지 주요 가정들이 내재되어 있는 고정된 구조 또는 시스템에 그 기반을 두고 있다.

품목구조는 4개의 수준으로 분류되어 있으며, 첫째 음식 및 음료, 운수, 그리고 의료서비스 등 대분류(major group)로 시작된다. 7개의 주요 그룹들은 69개의 지출군(expenditure class, EC's)으로 이루어져 있으며, 신선한 과일(EC11)과 병원 및 기타 관련서비스(EC 57) 등이 이에 속한다. 지출군은 다시 품목층(item strata)으로 지칭되는 207개 그룹으로 분리되며, 이는 지표가 구성되는 가장 하위수준에 해당한다. 품목층의 예를 2가지 들어보면 사과, 그리고 간호 및 재택 요양서비스(convalescent home care)이다. 품목범주는 상호 배타적이고 모든 소비자지출을 포괄하지만, 새로운 상품 및 서비스가 기준시점에 판매되지 않은 경우에도 자동으로 표본에 추가되는 것은 아님을 인지하는 것이 중요하다. 이는 단지 모든 상품 및 서비스가 도입되었을 경우 새로운 품목층을 만들 필요가 없다는 것을 의미한다. (이는 부분적으로는 “기타”로 불리우는 수많은 품목범주들을 통해 가능하게 되었다.)

각 상품층 내에는, ELIs로 약칭하는, 기초품목(entry level items)들이 정의된다. 지표들을 이 수준에서 구성하지는 않는다. 사과와 예와 같이 여러 개의 품목층이 하나의 ELI만을 갖고 있다. ELIs는 품목의 최하위 표본추출단위이다. 이는 자료수집가들이 각 표본상점에서 품목 표본추출을 시작하는 수준의 품목 정의에 해당한다. 예를 들어, West Terre Haute Indiana에 위치한 “Y” Foods, Inc, 기업에서 판매하고 있는, 모델 41303 41082, 4 3/10인치 길이에 플라스틱 용기가 포함된 브랜드 “X” 유아용 체온계에 대한 가격들이, 의료서비스 등 대분류, 의료상품 구성요소, 처방전 없는 의약품(EC 55), 처방전 없는 의약 및 의료장비(품목층 5503), 일반사용을 위한 의료장비(ELI 55032)를 위해 수집될 수 있다.

상점 표본은 지역 단위로도 추출되고 있다. 어떤 상점이 선정될 확률은 해당 품목범주의 조사지역 총지출에서 그 상점이 차지하고 있는 비중을 비례한다. 이는 선정된 품목의 가격을 소비자 구매장소를 대표하는 곳에서 측정하기 위한 것이다. 이는 또한 상점이 구매의 중요한 측면인 동시에 가격변화를 구성하는 요소인 것으로 가정하기 때문이다. 이 가정에 의하면 동일품목의 가격이 상점별로 다른 것은 구매에 있어서 서비스 품질 또는 위치상 편의 등의 측면에서 차이가 존재함을 의미한다. 그리고 소비자는 이와 같은 구매의 측면들에 대해 시간에 걸쳐 동일한 비율로 차액을 지불할 것이다. 이 가정이 성립하지 않는 경우, 예를 들어 일부 상점들이 다른 상점들에 비해 높은 성장률을 보일 때, 이 방법론을 사용할 경우 2가지 문제점이 발생한다: 즉, 보다 많은 가격측정치들 획득하거나 또는 보다 선호도가 높은 상점에 높은 중요성을 부여하는 것이 불가능하며, 동일한 품목을 서로 다른 장소에서 구입할 경우 직접 품질차이를 비교할 수도 없다.

품목 및 상점 표본을 “새롭게(refresh)” 하기 위한 조치가 있는데, 이는 표본교체로 지칭되며, 각 품목의 표본선정 확률을 최근 소비지출에서 차지하고 있는 그 품목 비중에 비례하도록 특정품목들의 표본을 추출해낸다. 매년 표본 중 대략 20퍼센트가 교체되므로, 전체적인 교체는 5년이 소요된다. 교체되어 들어온 품목들과, 이로 인해 교체된 품목들은 직접 비교되지 않는다. 이 조치는 교체시점 당시, 원래 품목과 교체되어 들어온 품목간 품질로 조정된 가격이 동일한 것으로 가정한다.

노동통계국 절차는 이전 가격조사 품목이 매진, 판매중단, 또는 영구히 판매 중지될 경우 이를 대신하는 가격조사 품목 선정을 고려하고 있다. 조사원들에게는 동일 ELI내에서 품목을 대체 또는 교체할 경우 사용할 지침이 주어지며, 그 내역의 비교가능성에 대한 판단이 내려진다.(그러나, 교체된 상품이 반드시 사라진 상품으로부터 시장점유를 빼앗은 것으로 하도록 의무화하는 규정은 없다.) 교체된 상품이 비교 불가능한 것으로 판명되면, 노동통계국은 대부분의 경우, 유사품목의 전 시점에 비한 가격변동을 제거한 가격 격차가 품질 격차로 인해 설명되는 것으로 가정한다. 어떤 경우에는, 품질격차 측정이 시도된다. 대체품 가격조사의 메커니즘을 촉발시키는 것은 한 품목이 소멸되는 것이라는 사실에 주의하라.

이미 수립된 품목층(stratum)에 속하지 않는 신상품, 즉 기준시점 이후에 도입됨에 따라 참조 장바구니에 포함되지 않은 신상품에 대해 품목대체 및 표본교체시 특별한 우선 순위가 주어지지 않는다. 따라서 이들은 종종 다음 10년 주기 지수개편에 이르기까지 지수에 포함되지 않는다.(더욱이, 과거 CPI는 개정되지 않기 때문에, 신상품이 미치는 파장은 소급 측정되지 않는다.) 도입 후 수년간 지수에 포함되지 않은 중요한 신상품 중 흔히 인용되는 사례는 에어컨디셔너와 VCR 등이다. 핸드폰은 CPI에 1998년 개정시 포함될 것이다.

이 방법론이 장바구니가 개정되기 이전에는 신상품 도입을 확실히 하지 않는 한편, 장바구니 개정(update)빈도를 높임으로써 시의성을 개선하는 일은 문제의 일부분만을 해결한다. 신상품의 품질을 경쟁상품과 직접 비교하는 일은 곤란한 경우가 많다. 더욱이, 신상품의 파장을 적절히 산출하기 위해서는 흔히 다른 품목군의 상품들과 비교를 거쳐야 한다. 현행 품목구조로 인해 CPI가 의술을 대체하는 의약품, 신문을 대체하는 전자정보 서비스, 자동차 구입을 대체하는 리스, 극장관람을 대체하는 비디오 임대 등을 완전히 포착하는 것이 저해된다. 시간이 지남에 따라 성공을 거둔 상품들의 가격변동은 CPI에서 보다 높은 비중을 갖게 된다. 그러나, 밀접한 대체관계에 있는 상품들이 서로 다른 품목군에 속해있을 경우 여러 품목군에 걸쳐 가격 파장을 완전히 측정하는 것은 불가능하다. 품목구조의 목적이 여러 가지 존재함에도 불구하고, 가장 중요한 목적은 지수추정에 있다.

특정 위치에 있는 특정 상점에서 측정된 특정 상품 및 서비스 가격은 통합되어 품목군-지역별 지수(item group-area indexes)를 형성하고, 이 지수는 품목분류구조의 상위 분류방향으로, 또는 지리적 구역을 기준으로 가중 평균되어, 전국 CPI를 형성한다. 가중치들은 구매시점조사(POPS), 소비지출조사(CES, 이는 그 세부사항에 있어서 그다지 풍부하지 못하다), 그리고 선정된 상점에서 특정 상품 또는 서비스를 개시하는 데에 사용되는 통계적 접근방법으로부터 도출된다. 이 설계는 시간이 지남에 따른 수량변동 의 수집을 고려하고 있지 않다(장바구니를 불변으로 가정하기 때문에, 이는 CPI 구성에 불필요하다).

산술평균을 이용하여 서로 다른 종류의 사과와 같은, 품목군 내의 가격변동을 통합하는 것은 가격변동시 수량 가중치를 불변으로 둔다는 제약조건 하에서 가능하다. 다음 부분에서 논의하는 바와 같이, 산술평균은 일정한 상식적 요건을 충족시키는 데에 실패한다.

이 방식으로 가격이 통합된 품목간 대체가능성이 높을수록, 이에 따르는 지수의 대체바이어스 또한 높다. 수량변동이 없다는 가정을 대신하여 지출비중에 변동이 없다는 가정을 사용할 수 있다. 이는 기하평균을 통해 가능하며, 이 경우 가격상승은 이에 비례한 수량감소로 상쇄된다. 예를 들어, 만약 granny smith 사과의 가격상승 10%가 구입된 수량감소 10%와 연결된다면, 기하평균이 시장반응을 포착하기 위한 적절한 방법이 될 것이다. 만약 가격상승과 연결된 수량변동이 없을 때에는 산술평균이 적절할 것이다. granny smith 사과의 경우, 여타 종류의 사과가 판매됨으로 인해 시장반응이 가격상승을 상쇄하고도 남는 규모로 발생할 수 있다. 이 현상이 발생할 경우, 기하평균을 이용하는 것은 시장반응을 과소 평가한다.

지리적 지표들의 공표가 여러 지역에 걸친 가격수준이 아니라, 오히려 CPI 변동률의 비교를 가능하게 하는 것임을 지적해둘 가치가 있다. 만약 변동률이 서로 다르다면, 필연적으로 어느 시점에서 가격 수준 또한 서로 다를 것임이 명백하다. 과연, 가격수준의 차이는 여러 지리 구역에 걸친 생계비 비교로, 현저한 관심대상이 될 것이다. 그럼에도 불구하고 이 방법론은 이 비교자료를 제공하지 않는다. 지리 구역은 표본설계시 중요한 역할을 수행한다. 그러나, 현재 구성되고 있는 방식의 지리적 지역지표들은 전국 CPI로 향하기 위한 통합의 한 단계 이외의 목적을 충족시키지 못한다.

요약하면, 표본추출 기법들을 이용하여 어떠한 품목의 가격을 어떠한 상점에서 조사하는지 결정된다. 이 방법론은 그 설계상, 역사적으로 대략 10년에 한 번씩 장바구니가 업데이트되기 이전에는 품목범주별 가중치를 변동시키지 않는, 고정 장바구니 개념에 충실하도록 요구한다. 이에 따라 이 방법론은 일부 신상품을 포착하는 데에 실패한다.

또한 그 가정상 소비자들이 보다 낮은 가격에 반응함에 따라 일부 상점이 여타 상점들에 비해 빠르게 성장하는 현상을 포착하지 못함으로 인해, 서로 다른 상점에서의 구매 경험을 직접 비교하지 못한다. 가격 지수가 구성되는 수준 중 가장 세부적인 것은 각 지리적 구역에 대한 207개 품목군이다. 지리 구역 물가지수는 전국 CPI를 향하는 과정에서 특정지역의 가격변동 추정치를 제공하기 위해 구성되나, 지역간 생계비 비교자료는 제공하지 않고 있다. 가격지수들은 전국CPI에 이르기까지 순차적으로 보다 넓은 범주로 통합된다.

결론적으로, 생계비 척도로서 CPI를 개선하기 위해서는 결정적인 가정들을 재검토하는 것부터 시작해서, 자료원천의 최적화기준 조정, 그리고 라스파이레스 지수 산식의 탈피 등 일정 범위의 이슈들을 다루는 일이 필요하다. 위원회의 권고사항들은 제Ⅷ부에서 소개된다.

#### **Ⅳ. 소비자물가지수와 생계비 지수 : 측정 이슈**

생계비 지수는 2개의 서로 다른 물가집합을 통하여 동일한 후생수준(복지, 효용, 생활수준 등으로도 알려져 있음)을 달성하기 위한 최소한의 지출을 비교하는 것이다. 대부분의 경우 이는 2개의 시점간 비교로 간주된다. 이론을 지수 숫자 생산에 실제로 적용하는 다른 어떤 경우와도 마찬가지로, 생계비 지수를 추정하기 위해서는 가정들, 방법론, 자료수집 절차 그리고 지수 숫자구성 등이 필요하다.

CPI에는 두 개의 잠재적 바이어스 집합이 존재한다: “이상적인” 생계비 지수에 대비한 바이어스들과 참조시기 내에서(within its own terms

of reference) 발생하는 바이어스들이다. CPI의 강점은 개념에 내재되어 있는 단순성에 있다: 여러 시기에 걸쳐 상품 및 서비스에 대한 고정된 (그러나 대표성 있는) 장바구니의 가격을 조사하는 것이다. 그 약점 또한 동일한 개념으로부터 파생된다: “고정된 장바구니”는 소비자들이 가격변동과 새로운 선택기회에 반응함에 따라 시간이 지나면서 그 대표성을 점차 잃어간다.

소비자들은 값비싼 상품들로부터 상대적으로 가격이 하락한 상품들로 대체함으로써 가격변동에 대응한다. 세상이 변화함에 따라, 이들은 쇼핑 상점, 다양성, 그리고 전적으로 새로운 상품 및 서비스에 대한 선택기회에 직면하게 되고, 이에 대해서도 반응을 한다. 이 변경들은 이전의 “고정바구니”가 갖는 관련성을 더욱 약화시킨다(irrelevant).

빠른 속도로 변화하는 세상에서 그 관념에 충실하려고 노력하는 와중에, 현행 CPI 절차들은 어려움에 직면한다. 할인점의 등장과 같은 변화를 무시하거나, 표본으로부터 품목들이 교체되어 나가고 새로운 품목들로 대체되는 등의 변화에 어떤 방식으로 대응하려고 할 경우 바이어스가 발생한다. 다소 짜깁기에 가까운 구조 속에서 “고정바구니”의 가격조사를 흉내내는 방법으로 변화를 포착하고자 하는 것은 필연적으로 결과되는 지수에 다른 문제점들을 야기한다. 서로 다른 이 바이어스들은 겹치고 있으며, 대체 바이어스(substitution bias); 산식 바이어스(formula bias); 상점대체 바이어스(outlet substitution bias); 품질 변화; 그리고 신상품 바이어스 등 몇 가지 표제 하에서 논의된 바 있다.

“순수한” 대체바이어스가 설명하기에 가장 용이하다. 애초의 기준시점 1과 이어지는 시기 2를 서로 비교하는, 매우 정형화된 예를 살펴보자. 단순화를 위해, 소고기와 닭고기 2개의 상품만 존재하는 가상적 상황

을 그려보자. 시기1에서, 소고기와 닭고기 1파운드당 가격은 \$1로 동일하며, 소비되는 수량 또한 1파운드로 동일하다. 따라서 총지출은 \$2이다. 시기2에서, 소고기는 닭고기에 비해 2배로 비싸다(1파운드당 \$1.60 vs \$0.80). 그리고 소비자들이 소고기를 상대적으로 덜 비싼 닭고기로 대체함에 따라, 소고기(0.8파운드)에 비해 훨씬 많은 닭고기(2파운드)가 소비된다. 시기2의 총지출은 \$2.88이다. 해당 자료는 표1에 제시되어 있다. 우리는 이 두 개의 상황을 어떻게 비교할 수 있는가? 사실 2가지 방법이 있는데, 각각 약간씩 다른 질문을 하고, 따라서, 별로 놀랍지 않게, 다른 해답을 제시하고 있다.

표 1 : 대체바이어스의 가상적인 예								
					상대가격		상대비중	
	시기1 가격	시기1 수량	시기2 가격	시기2 수량	P2/P1	P1/P2	1	2
소고기	1	1	1.6	0.8	1.6	0.63	0.5	0.43
닭고기	1	1	0.8	2.0	0.8	1.25	0.5	0.57

가장 단순한 비교는 “현재 나의 상황(시기2)에서, 애초(시기1)에 구입했던 것과 동일한 수량을 구입하기 위해 얼마나 더 지출해야 하는가?”를 질문하는 것이다. 이는 CPI에서 제기하는 질문이다. 시기1에 대한 시기2의 상대 가격지수는 애초 시기1의 소비 바구니를 계산과정에서 가중치로 사용한다. 소고기 1파운드와 닭고기 1파운드의 가격은 시기2에서 \$2.40이다. 시기1에 대한 시기2의 가격지수는 1.20(2.40/2.00), 즉 20% 증가이다.

직관적으로, 이 계산방식이 진정한 생계비변동을 측정하는 데에 있어 상향(대체)바이어스를 보이는 이유는 쉽게 이해할 수 있다. 이는 소비자가 소고기를 (보다 값싼)닭고기로 대체하지 않는 것으로 가정한다.

현실세계에서는, 가상적인 예에서와 같이, 소비자들이 상대가격 변동에 대응하여 지출패턴을 변화시키고, 따라서 부분적으로는 자신을 가격 동요로부터 분리한다.

이를 대신하는 접근방법은 “동일한 상품 및 서비스에 대해, 현재 상황(시기2)에서 나는 애초 시기1에 지배적이었던 가격하에 지출했을 금액에 비해 얼마나 많이 지출하고 있는가?” 질문을 던지는 것이다. 이 가격지수는 시기2의 지출(\$2.88)을, 현재(시기2)의 장바구니를 애초 가격하에서 구입할 경우 드는 비용(소고기 \$0.80와 닭고기 \$2.00를 합하여 \$2.80)과 비교한다. 이 물가지수는 1.03이며, 이는 단지 3% 증가에 해당한다. 이 접근방법은 상품대체를 과대 평가함으로써 진정한 생계비 상승을 과소 평가한다.

생계비 지수에 내재되어 있는 관점은 상대가격 변동에 따른 상품대체를 반영한다는 것이다. 생계비 지수는 “소비자가 이어지는 시기(시기2)에 동일한 후생수준을 누릴 수 있도록 하기 위해서는 애초(시기1)의 지출을 얼마나 증가(또는 감소)시킬 필요가 있겠는가?” 질문에 대한 답을 제시한다. 이 질문에 대한 해답을 얻기 위해서는 일견 소비자의 선호에 대한 구체적인 정보가 필요한 것으로 보일 수도 있지만, CPI에서 채택하고 있는 기존의 고정 가중치를 대신하여 “최고(superlative)” 지수 산식을 사용함으로써 훌륭한 근사치를 얻을 수 있다.

『최고 지수』 숫자의 개념은, 미국 경제학자, Irving Fisher(1922)에 의해 소개되었다. 이는 타당성 있는 일정 기준들을 충족시키고 따라서 아래 제시된 그의 “이상적” 지수(ideal index)에 밀접히 호응하는 지수 숫자들을 설명하기 위한 것이었다. 이 개념은 캐나다 경제학자, Erwin Diewert(1976)에 의해 일반화되었으며, 생계비 지수와 같이, 내재된 경

제지표에 대한 만족스러운 근사치를 제공하는 모든 지수 숫자 산식을 설명하는 데에 이용되었다. CPI는 고정가중치 산식에 바탕을 두고 있는데, 이는 이러한 근사치를 제공하지 못하고, 타당성 있는 이 기준들을 충족시키지 못하며, 더욱이 상향편중을 지닌 것으로 알려져 있다. 최고 지수는 고정가중치 지수와 동일한 상품 및 수량정보를 요구하나, 두 기간 중 하나를 기준시점으로 처리하기보다는 두 기간 사이를 중재(interpolating)하는 것을 그 내용으로 한다. 이를 달성하는 데에는 두 가지 방법이 있다.

시기간 중재를 위한 첫 번째 접근방법은 두 가지 고정가중치 지수 - 애초시기와 이어지는 시기를 각각 “기준”시기로 사용한 것 - 의 기하평균을 이용하는 것이다. 기하평균은 두 지수의 곱에 대해 제곱근을 취한 것이다. 이는 Irving Fisher(1922)가 최초로 고안하고, 이에 대한 경의로 현재 “Fisher 이상적 지수”로 지칭되는 이상적인 지수이다. 우리의 예에서, 이는 1.11, 즉 11% 증가에 해당한다. 비교할 경우, 시기1을 “기준”시점으로 삼고 있는 CPI-형 고정가중치 지수는, 9% 상향 바이어스를 지니고 있다(1.20에서 1.11를 차감). 이에 대해, 시기2를 “기준”시점으로 하는 고정가중치 지수는 8% 하향 바이어스를 지니고 있다(1.03에서 1.11를 차감). 경제분석국은 Fisher 이상적 지수를 미국 국민소득 및 생산계정에 관한 자료축적에 채택하였다.

중재를 대신하여 사용할 수 있는 방법은, 두 기간의 가중치 평균과 동일한 상대 비중을 부여한 후 가격상승률의 가중평균을 사용하는 접근 방법이다. 이는 그 고안자 중 한 명인 핀란드 통계학자 Leo Tornqvist를 기념하여 “Tornqvist” 지수로 지칭된다. 우리의 예에서 이는 1.10, 즉 10% 증가에 해당한다. 우리는 두 개의 최고 지수 산식으로부터 도출된 생계비 지수 근사치가 매우 유사한 것으로 결론짓는다. 두 가지

고정가중치 지수의 바이어스 추정치 또한 유사하다. 노동통계국은 생계비 변동의 척도로서 고정가중치 지수가 갖는 바이어스를 평가하기 위해, 고정가중치 지수를 Fisher 이상지수 및 Tornqvist 지수들과 비교한 바 있다.

CPI의 대체 바이어스는 그 규모가 얼마나 되는가? 이 질문의 답을 얻으려면 위에서 설명한 CPI 구조의 계층제적 성격을 감안해야 한다. 애초에는 207개 품목군과 44개 지역 지수들이 위해 통합되어 CPI를 형성하는 과정에서 발생하는 상위 대체바이어스(Upper Level Substitution Bias)에 초점을 맞추는 것이 유용하다. 노동통계국은 이 목적을 위해 고정 가중치를 이용하며(가중치는 소비지출조사(CES), 즉 가구 지출패턴에 대한 조사로부터 도출된다), 따라서 닭고기의 소고기 대체, 사과와 오렌지 대체 등을 무시한다. 노동통계국은 고정가중치 지수와, 우리가 설명한 중재방법 중 하나를 통해 도출한 지수와 비교함으로써 이 형태의 대체바이어스를 측정하였다. 그 추정치는 제VI부에 제시되어 있다.

CPI 대체바이어스 중 두 번째 종류는 하위 대체바이어스(Lower Level Substitution Bias)인데, 대략 71,000개에 이르는 상품 및 서비스 가격과 주거비용에 대한 정보가 207개 품목 및 44개 지역을 위한 지수로 구성될 경우 발생한다. 지수구성 중 이 부분은 CES 및 소매상점의 POPS로부터 산출된 확률을 이용한 확률표본을 그 내용으로 한다. 이는 개별품목이 특정 소매상점에서 구입될 가능성을 반영하기 위한 것이다. 이를 지출비중의 역할을 확률이 수행하는 또 하나의 고정가중치 지수로 받아들이는 것이 유리하다. 1978년 이래로 하위수준의 지수구성 산식은 상위수준의 고정가중치 지수와 매우 흡사했다. 산술평균을 이용하여 가격변동을 합산하는 것은 최하위 수준의 묶음에서 특정 상

품 또는 서비스의 상대가격이 변화할 경우 대체가 일어나지 않는 것으로 가정하는 것이다. 예를 들어, granny smith 사과가격이 상승할 때, 이는 예를 들어, delicious 사과로의 대체가 없는 것으로 가정한다는 것이다. 노동통계국은 이 고정가중치 지수와 하위 수준 가격에 대한 기하평균을 비교함으로써 하급 대체바이어스를 측정하였다.

지수구성의 하위수준에서 고정가중치 지수의 주요 문제점 중 하나는 시간 역전의 불가능성(failure of time reversibility)이다. 지수에 대한 이 단순하고도 직관적인 요구사항 또는 ‘시험’은, 내재된 가격들이 서로 역전될 경우 지수가 계속 동일해야 한다는 것이다. 예를 들어, 표1의 소고기 가격이 시기1에 1.0에서부터 시기2에 1.6으로 상승하나, 시기3에 다시 1.0으로 하락함으로써, 시기1과 2 사이 변화를 되돌려놓는 것으로 가정하자. 고정가중치 지수는 시기1과 2 사이에 60% 상승하지만, 시기2와 3사이 단지 37.5% 만 하락한다. 따라서 시기1과 3 사이 “소고기”지수가, 시간 역전이 가능하기 위해 요구되는 바와 같이 0이 아니라, 22.5% 또는 기간당 11.25%의 상승을 보이게 된다.

기하평균은 시간역전 시험을 충족시킨다. 이는 기간간 가격비율의 곱에 대한 제곱근을 바탕으로 하기 때문이다. 표1의 소고기가격 예를 보면, 시기1과 2 사이 가격비율은 1.6인 데에 비해, 시기2와 3 사이 가격비율은 0.625이다. 이 두 개의 가격비율을 곱하면 1인데, 이는 시간역전을 위해 요구되는 바와 같다. 따라서 기간 당 평균 가격상승률은 영이 된다. 시간역전 의 속성은, 기하평균이 지수 숫자 구성의 하위 수준에서 여러 접근방법의 비교기준으로 광범위하게 이용되는 계기가 되었다. 예를 들어, Moulton과 Smedley(1995)는 이 수준에서 노동통계국의 고정가중치 접근을 가중기하학적 접근과 비교한 바 있다.

Diewert(1995)는 하위수준의 지수숫자 구성을 위한, 서로 대체 가능한 접근들의 속성을 세부적으로 검토하였다. 이는 하위수준의 지수 숫자를 위한 여타 합당한 요구사항들뿐 아니라, 시간의 역전도 포함하고 있다. Shapiro와 Wilcox(1996b)는 시간에 걸친 상대가격의 공분산성에 기초를 둔 기하학적 접근에 대해 우아한 합리화근거를 제공한 바 있다. 이 공분산성이 작다는 전제하에, 수정된 기하평균은 대체로 내재된 생계비지수에 대해 불편성(unbiased)을 지니고 있다. 그리고 불편성 결론을 도출(characterization)하기 위해 내재된 소비자 선호체계에 대한 정보가 반드시 요구되는 것은 아니다.

수정 기하평균은 하위수준에서 지수 숫자구성 방법을 평가하기 위한 기준으로 널리 이용되어 왔다. Diewert(1995)가 수행한 실증연구 검토는 유용하다. Moulton과 Smedley(1995)의 작업에 더하여, Carruthers, Sellwood, 그리고 Ward(1980)가 영국에서 이와 같은 종류의 연구를 수행하였고, Schultz(1994)는 캐나다에서, Dalen(1994)은 스웨덴에서, 그리고 Woolford(1994)는 호주에서 이를 수행하였다. 이 연구들은 노동통계국에서 사용되는 것과 같은 고정가중치 지수들이, 상향 바이어스를 지니고 있음을 보여준다; 편중의 크기순서는 미국에서 Moulton과 Smedley(1995)가 수행한 연구에서 암시된 것과 유사하다. 이 문제들은 캐나다통계청과 같이, 점점 증가하는 수의 통계기관들이, Irving Fisher(1922)의 조언에 따라 산술평균을 버리고, 이를 기하평균으로 대체하도록 유도한 바 있다.

하위 수준에서 고정가중치 지수를 적용(implement)하는 데에는 어떤 상대적으로 미묘한 문제가 발생하였다. 표본 품목들이 이전의 가격관 측치가 제공되어 있지 않은 품목들로 대체될 경우, 이 공백을 메우기 위해 대체품에 기준시점 가격을 자동배분(impute)해야 한다. 노동통계

국에서 이를 위해 채택한 절차들은 CPI에서 이용된 가격과 대체상품의 가중치를 연계시키는 결과를 초래했으며, 하급 대체바이어스의 중요한 구성요소인 바이어스를 발생시켰다. 이 문제는 70,000개의 상품 및 서비스 가격표본과 주거비용 표본에 포함된 품목들을 교체하는 동안에도 발생한다.

전체 하급 대체바이어스 추정치는 고정가중치 지수와 가중 기하평균 간 차이를 통해 산출되는데, 고정가중치 지수는 노동통계국에서 1978년 소개한 가격자동배분 방법에 기초를 두고 있다. 1995년과 1996년 노동통계국은 가격추정치의 “평활(seasoning)”에 바탕을 둔 새로운 절차들을 도입하였다. 평활은 가격 자동배분시점과 품목이 실제 CPI에 도입된 시점 사이 기간을 연장하는 것이다. 이 기간을 연장함으로써 개별품목의 가중치와 가격 사이 연결고리를 끊고, 바이어스를 감소시키는 것이 가능하다. 그러나, 고정가중치 산식에 관련된 바이어스는 존속한다.

중간보고서(Interim Report)는 우리가 “산식 바이어스(formula bias)”로 지칭했으며, 현재는 하급 대체바이어스로 지칭되고 있는 것이 노동통계국에 의해 제거되기를 기대했었다. 노동통계국은 적절한 곳에 “평활”을 도입함으로써 그 절차를 변경했다; 이는 가격 자동배분 방법들에 기인한 바이어스를 제거한 반면, 하위수준에서 고정가중치 산식을 이용한 데에 기인한 바이어스에는 영향을 미치지 못하였다. 따라서, 우리는 이하에서 노동통계국이 고정가중치 산식을 가중 기하평균으로 대체하도록 권장하였다. 이는 Moulton · Smedley(1995)의 중요한 연구에서 실험끝에 시행 가능한 것으로 판명되었다. 이 연구의 대상기간 연장을 위한 추가작업이 진행중이다.

노동통계국이 하위수준의 합계에서 고정가중치 지수를 도입한 것은 그 당시 합계의 상위 및 하위수준 모두에 대해 지수화의 일관성을 가져다준 것으로 받아들여졌다. 그러나, 고정가중치 접근방법이 상위 수준에서 갖는 단점들은 하위수준으로 전달된다. 최고 지수 산식은 상위 수준에서 내재된 생계비 지수의 만족스러운 근사치를 제공해야 한다. 이는 CPI에서 채택한 고정가중치 지수 산식과 관련된 바이어스를 방지한다. 유사하게, 하위수준의 수량 및 지출관련 정보를 결여한 상태에서는, 가중 기하평균 산식에서 내재된 생계비지수에 대한 좋은 근사치를 얻을 수 있다.

소고기와 닭고기의 예와 같이, 상대가격 변화에 대응하여 소비자들이 구입상품을 변경하는 것처럼, 이들은 구매장소 또한 변경한다. 새로운 할인점의 개점은 소비자들에게 이전에 비해 저렴한 가격으로 구매할 기회를 줄 수 있다. 현재로서, CPI 절차는 소비자의 상점변경에 따르는 이 감소들을 무시한다. 그러나, 만약 소비자들이 오직 최대한 저렴한 가격으로 상품을 획득하는 데에만 관심이 있다면, 우리는 모든 상품들이 모든 상점에서 동일한 가격으로 판매되는 것을 관측할 것이다. 그 대신, 우리는 할인점과 회원제 창고점에서 저렴한 가격을 관측하는 동시에, 슈퍼마켓에서는 중간 가격, 그리고 편의점에서는 보다 높은 가격을 관측할 수 있다. 명백히, 소비자들은 가격에 관심 있을 뿐 아니라, 점원의 접근가능성, 포장서비스, 그리고 집과 여타 상점들간 거리 등 서비스의 수준에도 관심을 갖고 있다.

현행 CPI절차들은 소비자들이 상점을 전환할 때 발생하는 가격변동을 무시한다. 이로 인해 CPI는 상점간 가격격차가 전적으로 서비스 품질 차이를 나타낸다는 가정을 암묵적으로 포함하고 있다. 만약 경제가 교차되는 수준의 서비스 품질을 제공하는, 안정적인 상점의 집합으로 이

루어진 정적인 경제였으면 이 접근방법이 타당했을 것이다. 그러나, 보다 효율적인 유통기술로 인해 저렴한 가격의 상점이 확장한 반면, 보다 오래된, 가격이 높은 상점들은 축소되고 일부 폐업을 함에 따라 할인점의 시장점유율은 지속적으로 상승해왔다. 시장점유의 이 같은 전환은 수많은 소비자들이 가격차이에 반응을 하며, 가격격차가 서비스 품질의 차이로 완전히 상쇄되는 것으로 보지 않는다는 사실을 암시했다. 이와 같은 가격 하락을 CPI에 모두 반영함으로써 서비스 품질 격차를 모두 완전히 무시하는 것은 반대방향으로 착오를 일으킬 것이다. 진정한 가격변동을 서비스 품질변동으로부터 분리해내기 위해서는 추가적인 연구가 필요하다. 이 문제는 가격변동을 상품 품질의 변화로부터 분리해낼 필요성과 유사하다.

품질변화와 신상품은 측정에 있어서 가장 어려운 문제점들을 야기한다. 이는 신상품의 등장을 적시성 있게 포착하는 것; 신상품과 기존상품간 품질을 직접 비교하는 것; 신상품과 경쟁상품(서로 다른 범주에 속해 있는 경쟁상품, 예를 들어 새로운 의약품과 그것이 대체한 진료 행위)간 품질을 직접 비교하는 것; 또한 이 신상품들이 자신의 범주그룹 내외에서 여타 상품들을 추방하는 데에 따른 품질과 대체의 통합된 파장을 포착하는 것 등을 포함한다.

물가지수에 대한 한 유명한 전문가가 일반적인 이슈를 명백히 진술하였다: "...경제학에서 이질성은 거래와 관련되며, 단지 상품의 물리적 특성에 한정되지 않는다. 두 개의 거래가 외형적인 또는 암묵적인 속성의 서로 다른 묶음을 그 내용으로 하는 경우 언제나 이들은 질적으로 차이 난다. 판매, 판매와 함께 제공되는 서비스, ... 등의 측면에서 존재하는 차이점은 사안의 경제학적 측면에서, 우리가 일반적으로 "품질변화"라고 생각하는 물리적 변화와 정확히 일치한다"(Triplett(1990)).

예를 들어, 무엇이 언제(그리고 어떻게) 구입되었는가 뿐만 아니라, 언제 구입되었는지 또한 중요할 수 있다. 요일(time of week) 바이어스 또한 있을 수 있다. 노동통계국은 특정 품목 및 특정 종류의 상점들이 비정상적으로 할인판매를 하는 주말 또는 휴일에 가격을 수집하지 않는다. 주말 또는 휴일에 이루어지는 구매의 비중이 현저히 증가한 것으로 보이는데, 이는 맞벌이 가구가 더욱 우세해진 현상을 반영하는 것일 수도 있다. 우리가 아는 한 이 이슈가 체계적으로 연구된 바는 없으며, 우리는 노동통계국에서, 어쩌면 스캐너 자료를 포함시켜서, 이를 철저히 검토하는 데에 필요한 연구를 수행하도록 재촉하는 바이다.

이 이슈들을 완벽하게 다룰 경우 “평균적인” 또는 “대표적인” 소비자에 초점을 맞추는 문제가 더욱 강화된다. 서로 다른 소비자들은 서로 다른 취향과 시간 비용을 갖고 있으며, 따라서 새로운 상점 및 상품의 출현을 달리 평가한다. 이에 따라 일부(다수)는 슈퍼마켓에 의해 후생이 높아지고 다른 이들은 길모퉁이 구멍가게의 소멸에 따라 후생이 악화되는 것을 경험한다. CPI는 서로 다른 소비자들 또는 소비자 집단들이 갖는 독특한 특징들을 설명하기 위한 장치들을 갖추고 있지 못하다.

진정한 생계비 지수(COLI)를 얻기 위해 원칙적으로 적용될 여타 이슈들이 아직 있다. 범죄율 상승의 부정적인 영향과 이것이 수반하는 안전장치의 구입 및 보험 프리미엄 상승, 그리고 아이가 아플 경우 부모가 집에서 업무를 수행할 수 있도록 하는 정보기술 향상의 긍정적인 효과 등 2개의 예를 살펴보자. 이들이 “적어도 동등한 수준의 후생을 누리기 위해 필요한 최소한의 지출”을 산출하는 데에 포함될 것은 확실하다. 아래 제Ⅶ부는 이 문제 중 일부를 탐구한다.

## V. 품질 변화와 신상품

### 도입

품질 변화와 끊임없는 신상품의 출현으로 야기되는 곤란한 문제들은 “물가측정의 각개전투(house-to-house combat)”로 지칭되어 왔다. 이 장에서 우리는 신상품 바이어스를 품질변동 바이어스의 일부로 취급하고, 전체 바이어스 추정치를 별도의 품질변동 바이어스와 신상품 바이어스 기여분으로 분해하지 않을 것이다.

어떤 상품의 경우 품질변화가 빠른 속도로 진행되어온 반면 여타 상품에 대해서는 그렇지 못하다. CPI는 여타 상품에 비해 일부 상품의 품질변동 효과를 보다 성공적으로 포착한 바 있다. CPI는 일부 신상품을 다른 것에 비해 보다 신속하게 도입하였다. 품질변동 바이어스의 중대성은 상품군에 따라 너무도 큰 차이를 보이고 있다. 따라서 어떠한 경우에도 품질변동 바이어스의 중요성을 총체적으로 평가하는 일은 “상당히 세부적으로(down in the trenches)” 수행되어야 한다. 즉, 이는 소비지출의 개별 범주를 취한 후, 각 범주에 대한 품질변동 바이어스를 평가하고, 적절한 가중치를 사용하여 그 합계를 산출해야 하는 것이다.

분석을 더욱 복잡하게 하는 것은, 개별상품 수준에서 평가된 품질변동 바이어스가, 시간이 지남에 따라 현저히 변화한 것으로 보인다는 사실이다. 예를 들어, 노동통계국 방법론의 중대한 개선으로 인해 1960년대 중반 이전 CPI에 존재했던 신규 자동차의 상향 바이어스와 1980년대 중반 이후 의복의 하향 바이어스를 대체로 또는 전적으로 제거한 바 있다. 유사하게, CPI 임대료 지수에 있던 중요한 하향 바이어스의 원천이 1980년대 후반 제거되었다.

품질변동 바이어스에 대한 이전의 평가, 예를 들어 Shapiro와 Wilcox(1996c) 그리고 Lebow, Roberts, 그리고 Stockton(1994)은, 사전에 수행된 연구로부터 특정 상품, 예를 들어 가전제품(consumer appliances) 또는 자동차의 바이어스 추정치를 추출한 다음, 그 바이어스 추정치를 CPI에 있는 상품들의 가중치로 적용하고, CPI의 나머지 부분은 품질변동 비율이 영임을 가정한다. 이 접근방법이 바이어스를 중립적으로 평가할 가능성이 높은 것으로 보는 것은 아니다. 이는 단지 CPI의 나머지 부분이 불편성을 충족시킨다는 평가로 인해, 상대적으로 수행된 연구가 적은 CPI 구성요소들이 바이어스를 지니고 있는지 여부에 대한 답이 지극히 치우친 입장을 대변하고 있기 때문이다. 이들이 세심한 연구가 수행된 구성요소의 평균 바이어스 비율에 해당하거나 또는 전혀 바이어스가 없을 가능성은 일치한다. 이 장에서 우리는 CPI를 각 요소별로 평가하고, 범주들이 서로 연관된 것으로 보일 경우, 한 개 범주의 바이어스에 관한 연구를 다른 범주로 유추한다. 그럼에도 불구하고, 우리는 심층연구를 수행한 범주들과 사뭇 다른 것으로 판단되거나 또는 미측정 품질변동과 신상품이 상대적으로 중요하지 않았던 몇 개 범주에 영의 추정치를 부여한다.

품질변동과 신상품에 기인한 바이어스 문제가 위에서 언급한 여타 형태의 바이어스와 대체로 분리 가능한 한편, 이것이 전적으로 가능한 것은 아니다. 다른 연구들, 예를 들어 Gordon(1990)이 개발한 품질변동 바이어스에 대한 증거는, 우편주문 카탈로그와 소비보고서(Consumer Reports) 등의 자료원천을 사용하여 노동통계국 가격 관측치와 독립된 원천으로부터 직접 가격을 측정하기 위한 시도에 바탕을 두고 있다. 그러나, 동일한 상품에 있어서, 이 독립된 지수들과 CPI 사이 어떠한 차이점도, 품질변화와 신상품 바이어스뿐만 아니라, 기존의 대체바이어스(왜냐하면 상품 및 모델조합이 CPI에 비해 이를 대신하는 원천에

서 더욱 빠른 속도로 전환되기 때문), 상점대체 바이어스(왜냐하면 대체적인 물가 관측치는 종종 계속 변화하는 할인점 조합에 따라 조정되는 시장가격의 평균이기 때문), 그리고 산식 바이어스(왜냐하면 대체적인 지수들은 전에 논의한 산식 바이어스의 문제들로부터 자유롭기 때문) 등을 반영할 수 있다.

## 개념적 이슈

기존 상품의 품질변동과 신상품 등장으로 인해 야기된 어려움은, “두 시기간 동일한 후생수준을 달성하기 위해 요구되는 최소한의 지출비교”, 즉 생계비 지수의 정의로 되돌아옴으로써 가장 강조된다. 두 번째 시기에 공급되는 주어진 상품의 모델이 첫 번째 시기와 다른 품질 속성들을 구현하고 있을 때, “동일한 수준”은 무엇을 의미하는가? 그리고, 이보다 더 뿌리깊은 어려움으로, 첫 번째 시기에 공급되지 않았던, 전적으로 새로운 상품이 등장할 때, “동일한 수준”은 무엇을 의미하는가?

“상품주기”로 지칭되는, 점차 확산되고 있는(pervasive) 현상은 CPI의 신상품 바이어스를 평가하는 데에 결정적이며, 기존 상품의 신모델에 대해서도 마찬가지로 적용된다. 전형적인 신상품은 상대적으로 높은 가격에, 적은 판매규모로 도입된다. 곧 제조기술 개선과 판매증가로 인해 가격은 하락하고 품질은 향상된다. 예를 들어, VCR은 1970년대 말 \$1,000에, 어색한 전기 기계적(electromechanical) 제어장치와 함께 등장했다; 1980년대 중반쯤 가격은 \$200으로 떨어졌으며 전자 제어장치와, 광범위한 예약프로그램 기능을 갖추고 있었다. 상품주기의 후반에 이르러, 상품은 성숙기를 거쳐 결국 해당 상품군의 평균적인 상품에 비해 가격이 빠르게 상승할 것이다. 이 일련의 과정은 “U”-형태의

곡선으로 쉽게 형상화된다 - 소비자 장바구니에 대비해 주어진 상품의 가격은 높이 시작해서, 그 후 떨어지다가, 한동안 수평을 유지하고, 그리고는 다시 올라간다. CPI가 성숙기에 있는 상품들에 과도한 가중치를 부여하고 신상품에 과소 가중치를 부여하는 한, CPI는 상향 바이어스의 경향을 보일 것이다. 최근 몇 가지 학술연구, 가장 주목할 만한 것으로는 Berndt, Griliches와 Rosett(1993), 그리고 Berndt, Cockburn과 Griliches(1996) 등이 실제 판매되는 처방약 조합과 CPI에 포함되어 있는 더 오래된 제한적인 표본을 통해 서로 다른 물가지수를 산출한 것이며, 이 연구는 실제 판매되고 있는 모델과 신종(varieties)의 조합을 도입하는 일이 지연된다는 근거 하에, CPI에 현저한 상향 바이어스가 있음을 지적한다.

품질변화 및 신상품 바이어스를 평가하는 중요한 기준은 특정 모델과 상품의 시장점유율의 변천과정이다. 구모델에 비해 보다 값비싼 신모델이 등장할 때, 신모델이 시장점유를 획득할 경우, 우리는 신모델이 구 모델에 비해 양자간 가격차이보다 큰 정도로 품질이 우수하다는 결론을 내릴 수 있다.

동일한 기준은 상점대체 바이어스를 다루는 데에도 도움이 된다. 소비자들이 기존의 슈퍼마켓으로부터 개선된 상품구색과 다양한 농산물을 제공하는, 새롭고 보다 비싼 전문 식품매장으로 전환할 경우, 우리는 소비자 후생이 증진되었음을 유추할 수 있다. 지난 15년간 Wal-Mart가 더욱 저렴한 가격을 청구하고도 월등히 최고규모의 체인이 된 사실은, CPI에서 암묵적으로 가정하는 바와 같이, 소비자들이 Wal-Mart의 저렴한 가격을 열등한 서비스가 상쇄하는 것으로 생각하지 않고, 오히려 Wal-Mart가 기존에 제공되던 상점의 조합에 비해 가격과 서비스 면에서 보다 우수한 결합을 제공하는 것으로 본다는 사실을 의미한다.

7-11과 같은 편의점들이 높은 가격을 청구하고도 시장점유를 획득한 사실은 소비자들이 편의점들에 대해, 그들이 추가로 청구하는 가격보다 가치가 높은 추가적인 편리함을 제공하는 것으로 본다는 사실을 의미한다. 여러 소비자들은 Wal-Mart와 편의점에서 모두 쇼핑을 한다. 이들은 이전 상점들의 조합에 비해 특정 품목에 보다 높은 가격을 지불할 때도 있고 보다 낮은 가격을 지불할 때도 있다. 또한 시장점유의 전환은 새로운 조합이 개선을 가져온다는 사실을 시사한다. 외식업체 부문에서는, McDonalds와 같은 저렴한 가격의 패스트푸드 점포들, Olive Garden과 Red Lobster와 같은 중간 가격의 프랜차이즈, 그리고 일부 도심지역에서는 Tuscan, Thai 및 기타 외식전문만을 취급하는 값 비싼 고급 레스토랑 등으로 소비자들이 이동했는데, 이 경우도 동일한 평가를 내릴 수 있다. 의학 이미지장치(Trajtenberg, 1990)에서부터 아침 식사용 시리얼(Hausman, 1996)에 이르기까지 너무도 다양한 상품에 대한 중요한 일련의 학술연구에서 상품 다양성 증가에 상당한 가치를 부여한다. 따라서, “다양성의 가치”는 상점대체 바이어스와, 이 장에서 품질변동 및 신상품 바이어스를 평가하는 데에 양쪽에 걸쳐 모두 결정적인 중요성을 지니고 있다.

## 노동통계국 방법론

품질변동과 신상품 바이어스에 대한 우리의 논의는 CPI가 기존상품의 품질 변화를 처리하는 데에 사용하는 방법을 점검하는 데에서 시작한 후 신상품에 의해 발생하는 문제들을 다루게 된다. 노동통계국은 기존 상품의 모델 변화에 대처하기 위한 방법을 다섯 가지 보유하고 있다.

“직접적 비교”의 방법은 구 모델과 신 모델 사이 관측된 가격변동분을 모두 가격변화로 취급하며, 품질변화로 취급하는 부분이 없다. 품질은 향상될 수도 있고 악화될 수도 있기 때문에, 이 경우 필연적으로 바이어스가 존재하는 것은 아니다. 그러나 실제 대부분의 상품들은 지속적인 개선을 경험하는 경향이 있고, 가격변동 없이 보다 개선된 상품이 등장하는 경우도 흔하며, 이 경우 이 방법은 품질변동을 완전히 놓치게 된다.

“구모델 삭제” 방법은 구 모델과 신 모델 사이의 가격을 전혀 비교하지 않는다. 그 대신, 이 상품에 귀속 가능한 가중치는 동일상품군내 여타 상품의 평균 가격변화로 적용된다. 삭제방법을 이용하는 한, CPI는 상품주기 후반에 존재하는 일정한 품질의 상품들이 균형에 어긋날 정도로 많은 부분을 차지하게 된다.

“양모델의 연결” 방법은 신 모델과 구 모델이 동시에 판매되고 있을 때 사용될 수 있다. 이 경우 신 모델 등장 당시 두 모델간 가격격차는 두 모델간 품질격차의 가치에 대한 추정치로 사용될 수 있다. 위에서 지적한 바와 마찬가지로, 이는 신 모델이 시장점유를 확대해나갈 경우 품질변화의 과소평가를 초래할 수 있다. 또한, VCR의 가격은 지속적으로 하락한 반면 프로그래밍 기능과 재생 품질은 개선된 경우와 같이, 심지어는 신 모델이 구 모델에 비해 저렴하거나 구 모델과 가격이 동일한 경우에도 품질개선이 일어날 수 있다.

“신모델의 비용추정” 방법은 신 모델의 추가된 속성들에 대한 비용을 수립하려는 시도이다. 비용부과 방법과 관련된 현실적인 문제점들은, 사용빈도가 낮다는 점과, 다른 상품들에 비해 균형에 어긋날 정도로 자동차에만 적용되어 왔다는 점이다. 이는 비용부과 방법을 균등하지

못하게 적용하여 자동차에 비해 여타 상품들이 의심스러운 상대가격 상승 “압력(drift)”을 받을 가능성이 있음을 제기한다. 새로이 부각되고 있는 상향 바이어스의 원천은 자동차와 같은 상품들이 철강(더 이상 예전과 같이 녹슬지 않는다)과 타이어(훨씬 많은 마일동안 지속된다) 등 원자재의 품질개선의 혜택을 받는다는 사실이다. 자동차 생산과정의 이 투입요소 중 자신의 품질개선을 비용격차 이상으로 경험하는 것이 있는 한, 이는 노동통계국의 비용-기반 품질 추정절차에 포착되지 않을 것이다.

지금까지, CPI는 오직 의복 범주에 대해서만 품질변동의 가치추정을 위한 또 하나의 방법인, “hedonic 회귀분석 방법”을 도입했다. hedonic 접근방법은 품질변화 조정을 위해 제조비용 추정치를 대신할 수 있는 방법으로 볼 수 있다. 이 방법은 주어진 시기에 관측된 어떤 상품의 가격이 그 품질 속성들의 함수인 것으로 가정한다. 그리고 그 상품의 서로 다른 모델 가격들을 내재된 속성의 수량에 대해 회귀분석 함으로써 이 속성들의 자동배분(imputed)된 가격들을 추정한다. 따라서 hedonic 접근방법은 새로운 방법이라기보다는, 현실적 요인으로 인해 통상적인 방법에 비해 이 방법이 보다 적합한 경우 사용되는 비용추정을 대신하는 방법에 가깝다.

본질적으로 hedonic 지수를 위해서는 방대한 양의 자료가 필요하다. 모든 현대 산업사회에서 수천 개의 개별상품이 생산되는 것을 전제로 하면, 각 상품에 대해 완전한 횡단면 자료수집을 필요로 하는 hedonic 기법을 단번에 도입하는 데에는 상당히 무리가 따른다. 그러나 많은 경우 CPI 조사원들이 이미 수집한 자료를 hedonic 회귀분석에 사용할 수 있다; 이는 의복의 경우 이미 이루어지고 있다.

또 다른 가능성 있는 반대이유는, 몇 주 이내에 다시는 개정할 수 없는 월별 물가변동의 추정치 작성을 지향하는, CPI에서 요구되는 적시성 있는 방식으로 hedonic 지수를 구성하는 것이 불가능한 점이다. 그러나 이 주장은 역사적 자료를 바탕으로 공분산들을 추정할 수 있으며, 이전에 추정된 이 공분산들을 사용하여 새로운 상품이 등장했을 때 품질변동을 평가할 수 있다는 사실을 무시하고 있다. 이 접근방식은 TV와 PC에서 주로 볼 수 있는 바와 같이, 빠른 속도로 신 모델이 연이어 등장하는 상품범주에 특히 적합할 것이다.

노동통계국 방법의 이 목록은 적어도 4가지 상향 바이어스 원천을 드러내고 있다: 품질이슈를 아예 다루지 않는 직접비교 방법의 사용, 품질변화가 없고 상품주기의 후반에 있을 수 있는 모델을 바탕으로 가격변동을 측정하는 삭제방법의 사용, 모델간 품질향상이 가격격차에 비해 클 때 연결방법 사용, 그리고 원자재 및 투입요소 공급기업들이 품질향상을 달성하여 최종품 생산기업에 기여하는 것을 포착할 수 없는 비용방법의 사용등이다.

CPI가 상품의 가격상승 없이 소비자 후생증가를 가져오는 품질향상의 가치를 체계적으로 추정하기 위한 시도를 하지 않는다는 사실은 이보다 더 큰 어려움으로 작용할 수 있다. 예를 들어, 여러 개의 소비 전자제품과 가전제품은 수리횟수와 전력소모 감소를 달성했으며, 이러한 개선 중 CPI에 반영된 것은 거의 없다. 자동차(아래에 언급되어 있음), 가전제품, 그리고 기타 상품의 내구성 제고는 이와 유사한 바이어스 원천으로 작용한다.

## 신상품 바이어스

이제 신상품 바이어스의 이슈를 살펴보도록 한다. 상품주기의 진실성은 논란의 여지가 없으며, CPI가 많은 상품을 늦게 비로소 도입함으로써 상품주기 첫 단계에 전형적으로 일어나는 가격하락을 상당부분 놓치는 사실에 이의를 제기할 사람은 없을 것이다. 극단적인 예로 1951년에 이미 널리 판매되고 있었으나, 13년 후인 1964년에야 비로소 CPI에 도입된, 실내 에어컨디셔너를 들 수 있다. 보다 최근에는 전자레인지가 1978년, VCR과 PC가 1987년, 즉 처음 시장에 판매된 지 수년이 지난 후에 비로소 CPI에 도입되었다. 심지어 이보다 근래에 발생한 예로는, 현재 미국에서 사용중인 핸드폰이 36백만 개임에도 불구하고, 아직까지 CPI는 핸드폰 가격지수를 보유하고 있지 않다는 것이다. 따라서 자녀, 친지 또는 노령 부모의 행적을 추적하는 일이 보다 용이하게 된 데에 따른 소비자이익은 인플레이션, 실질생산, 또는 생산성에 대한 전국 척도에서 그 공적을 전혀 인정받지 못하고 있다. 심지어 더욱 최근에는, 미국 내에 핸드폰 가입자가 40백만을 넘었음에도 불구하고, 핸드폰은 아직 CPI에 도입되지 못한 단계이다.

신상품 바이어스의 두 번째 측면은 상품의 정의가 협소한 데에서 비롯된다. CPI에 새로운 상품이 드디어 도입될 때에도, 신상품 및 신상품과 동일 기능을 수행한 구상품 사이 가격 및 품질비교는 수행되지 않는다. 예를 들어 사람들은 때를 지어 비디오 테이프를 빌리고 있으나, 극장으로 가는 것에 비해 집에서 영화를 볼 때 드는 비용의 하락은 CPI에 반영되지 않는다. 유사하게, 전기타자기를 대체한 전자타자기, 그리고 그 후 워드프로세싱 및 맞춤법 체크 기능을 갖춘 PC, 또는 구식 단행본 버전에 비해 훨씬 적은 비용이 들고, 도서관을 여러 번 방문해야 하는 번거로움을 제거해준 CD-ROM 백과사전 등을 CPI에서는

고려하지 못했다. 그럼에도 불구하고 신상품이 이전 시기에 대응시킬 대상이 없는 진정으로 새로운 속성들을 구현하고 있는 경우가 많은 것은 불가피하다. 부모와, 집을 떠나있는 대학생 자녀간 새로운 유대 및 의사소통 집합을 제공하는 전자메일, 그리고 병든 자녀 또는 노령 부모와 사실상 끊임없는 접촉을 가능하게 하는 핸드폰 등은 단지 두 가지 예에 불과하다.

이러한 신상품 논의는 필연적으로 평균적 미국인의 생활수준 변동에 대한 보다 심오한 질문들로 귀결된다. 소비자 전자기술로 인해 가능하게 된 긍정적인 변화들을 범죄의 증가, 환경오염, 및 기타 “악화(bads)” 등에 대비해 경중을 가려야 한다. 이러한 이슈들은 아래의 제Ⅶ부에서 다루도록 한다.

## 상품 범주별 품질변화 및 신상품 바이어스

품질변동 바이어스는 규모, 방향 및 시기 측면에서 상품범주별로 차이가 난다. 따라서 품질변화 바이어스 규모에 관한 불확실성의 범위를 좁히기 위한 유일한 방법은 이용 가능한 증거를, 각 범주별로 살펴보는 것이다. 표2는 이 평가를 위한 지침제공을 위해 설계되었다. 왼쪽 열은 1995년 12월에 대해, CPI내의 주요 상품범주를 각각 “상대적 중요성”, 즉 퍼센트 비중과 함께 열거하고 있다. 이 장에서 우리는 각 범주별로, 품질변화 및 신상품에 관련된 바이어스에 대한 접근 가능한 증거를 검토한다.

일부 범주는 결론을 내릴 만한 자료가 거의 발표된 바 없다. 그러나 이 범주에서 품질변화 및 신상품에 기인한 바이어스 전체가 반드시 영인 것으로 가정하여 이전 연구를 추종하지는 않는다. 대신 우리는

‘두 기간간 동일한 후생수준 달성을 위해 필요한 최소지출의 비교’라는 생계비 정의의 맥락에서 바이어스가 취할 가능성이 높은 방향을 논의한다.

**1. 음식 및 음료.** Reinsdorf(1993)는 음식·음료 범주의 상향 바이어스에 대한 가장 극적인 증거를 제시했다. 그는 1980-1990 기간동안 50개의 좁게 정의된 품목에 대해 실제 지불된 연간 가격변동비율이 동일한 상품범주의 CPI수치에 비해 매년 2.0퍼센트 낮은 사실을 발견하였다. Reinsdorf는 그 당시에 이 차이가 상점대체바이어스를 나타내는 것으로 생각한 반면, 나중에는 오히려 산식 바이어스(formula bias)와 상점대체 바이어스가 혼합된 것으로 결론지었다. Reinsdorf의 연구에 대한 해석이 어떻든 간에, 품질변화의 증거를 나타내는 것은 아니다. 왜냐하면 그는 CPI에서 가격이 조사된 것과 동일한 상품들을 선택하였기 때문이다.

그의 연구 이외에, Hausman(1996)이 새로운 종류의 아침식사용 시리얼이 등장한 데에 따른 가치의 평가를 시도한 것을 제외하고는, 음식 범주에 대한 발표 증거가 거의 없다. 아마 새로운 종류의 포장상품보다 중요한 것은, 기술향상의 물결로 인해 겨울에 전형적인 슈퍼마켓에서 구입할 수 있는 신선한 과일 및 채소의 종류가 월등히 증가한 것과, 슈퍼마켓 제공서비스의 증가추세로 인해 작은 전문가게, 특히 신선한 생선시장과 갯 요리한 음식을 판매하는 조제식품 판매점들을 방문할 필요성이 없어진 것이다. 30년 전의 훨씬 제한된 선택가능성으로 제약 받는 대신, 오늘날 슈퍼마켓에서 구입 가능한 다양한 상품들로부터 선택할 수 있는 특권을 누리기 위해 소비자들은 얼마나 지불할 용의가 있을 것인가? 다양성과 편리함 향상의 가치에 대한 보수적인 추정치는 농산물을 제외한 가정 소비식품은 10%, 겨울시장의(뿐만 아니라

여름 농산물시장) 다양성 증가가 그렇게도 두드러진 농수산물은 20%, 그리고 수입맥주, 미니 양조장, 그리고 전세계로부터 들어온 수입포도주 유통이 월등히 개선됨으로써 생활수준 향상에 기여한 알콜음료는 5% 정도가 될 것이다. McDonalds에서부터 고급 레스토랑(위에서 논의한 바와 같이)에 이르는 모든 가격대의 외식의 다양성과 편리함 향상 또한 10%의 프리미엄을 부여할 수 있다. 표2의 연간 바이어스 비율은 이 추정 프리미엄을 지난 30년에 걸친 연간 기하 성장률로 전환함으로써 계산하였다.

2. 주거. CPI에서 가장 높은 비중을 차지하는 단일 요소는 단연코 주택이며, 그 내에서도 주거이다. 주거 구성요소는 1983년에 귀속임대료 접근방식으로 전환했으며, CPI-U-X1지수는 1967년에 이르도록 귀속임대료 개념을 소급 적용하여 주거비를 일관성 있게 처리하기 위한 시도이다. CPI 주거지표의 연간 변화율은 1967년에서부터 1983년에 이르기까지 CPI 주택임대지표를 매년 2.33퍼센트 초과한다. 그리고 이에 따라 공식CPI-U의 연간 변화율은 동일한 기간에 대해 CPI-U-X1를 매년 0.52퍼센트 초과한다. 노동통계국은 또한 1995년에 산식 바이어스 정정을 위해 방법론을 전환하였고, 1988년에는 연속되는 기간동안 점차 노후화하는 주거단위의 가격을 조사하는 데에 따른 “노후화 바이어스” 정정을 위해 방법론을 전환한 바 있다. Randolph(1988)는 1988년 이전 이 노후화 바이어스를 매년 0.3퍼센트로 추정하며, 이는 모든 유지 및 수리비를 제거한 순수 감가상각 효과를 나타내는 개념이다.

첫째, 우리는 Randolph의 노후화 바이어스가 과연 그 전체로서 바이어스로 취급되어야 하는지에 대해 회의적인 입장임을 밝혀둔다. 오래된 주택은 신규 주택에 비해 2가지 이유로 임대료가 낮다. 첫째, 수리 및 유지로 상쇄되는 정도에 비해 심각한 물리적 손상을 입을 수 있다. 그

러나, 둘째, 중앙 에어컨디셔닝과 같은 편의시설을 갖춘 보다 새로운 주택이 시장에 나오면서 가치를 잃을 수도 있다. 이와 같은 경제적 감가상각은 오래된 아파트에서 제공되는 서비스 품질저하를 나타내는 것이 아니라, 오히려 주거 편의시설에 대한 소득탄력성이 양의 값을 지니고, 국가의 1인당 소득이 증가함에 따라 사람들이 아파트 및 주택에 대해 보다 높은 품질을 기대하는 사실을 반영하고 있다. Gordon(1990)이 상세히 논의한 제트비행기의 등장을 통해 이에 대한 정확한 유추가 가능하다. 순수-제트 DC-8이 1958년 등장했을 당시, 프로펠러 동력을 이용한 DC-7 탑승의 품질은 하락하지 않았다. 오히려, 소비자들이 제트 비행기 탑승의 가치를 너무도 높이 평가한 나머지 DC-7의 수요가 사라진 것이다. DC-7은, 제트가 등장한 지 5년에서 10년내에 폐기되었으며, 이는 시기상조였던 것으로 평가된다. 장거리 항공여행이 프로펠러에서 제트 비행기로 전환된 데에 따른 이익 전체를 소비자들이 얻게 되었고, 오래된 아파트의 임대료 하락은 동일현상에 대한 덜 극적인 예를 나타내고 있다. 지금까지는 CPI 임대료조사에 포함된 아파트 품질변동이 거의 조사되지 않았다. “CPI 방법은 주거단위 횡단면표본의 임대료 변동을 시계열로 측정함으로써 대부분의 주거품을 일정하게 놓고 있다”(Randolph, 1988, p.359). 즉, 시간이 지나면서 주어진 주택의 임대료 변동이 계속 추적되며, 여타 주택들이 교체되어 들어가고, 겹치는 부분은 제거되는 방법으로 처리된다. 만약 일반적으로 최근에 건설된 단위일수록 더 많은, 그리고 더 나은 장치와 중앙 에어컨디셔닝, 그리고 지난 수십 년간 존재하지 않았던 여타 편의시설을 보유하는 경향이 있을 경우, 만약 소비자들이 이 편의시설을 추가 비용에 비해 높이 평가할 경우, CPI 임대지표가 상향 바이어스를 지닐 가능성이 있다. 가구들이 지속적으로 교외의 새로운 아파트단지와 Sunbelt 지역으로 이주하는 경향은 주택품질이 꾸준히 향상되는 과정의 일부일 수 있다. “시장점유” 시험은 수많은 가구들이 새로운

sunbelt 아파트를, 북부 및 동북부 도심중앙의 오래된 종류의 아파트에 비해 선호함을 시사한다.

미국 센서스 현행 주거보고서는 모든 임대 주거단위에 대해 매월 임대료 중앙값을 보고하고 있다. 1976년 대비 1993년 비율은 2.92(\$487/\$167)이다. 동일연도에 대한 CPI 임대지표 비율(산식 바이어스 또는 노후화 바이어스 조정을 거치지 않음)은 2.46이다. 여기에 내포된 CPI 성장률 격차는 매년 -1.00이다. 1973-1988의 또 다른 비교결과 매년 -1.10의 차이가 나타난다.

임대단위 품질에 관한 자료는 제한된 한편, 임대단위들이, 보다 많은 자료가 제공되어 있는 자가소유 주택단위와, 대략 비슷한 속도로 품질개선을 달성한 증거는 존재한다. 두 개의 핵심 척도에 따라 우리는 임대와 자가소유 단위들의 비교가능성을 확신하게 되었다(CPI는 임대와 자가소유 주거부문에 모두 임대지표를 사용하고 있으므로, 이와 같은 결론은 CPI에서 선택한 방법과 일맥상통한다). 첫째, 1970년과 1993년 사이에 사용중인 모든 단위(이 중 임대 단위는 대략 1/3에 해당)에서 평균 방수가 9.7% 증가한 반면, 임대단위에서는 7.8% 증가하였다. 이보다도 1인당 방수가 모든 단위에서 30.2% 증가한 것과 임대단위에서 27.0% 증가한 것이 아마 더 중요할 것이다. 이와 같은 비교집합은 임대 및 자가소유 단위에서 대략 비슷한 속도로 품질이 향상되었다는 견해와, 주택전체(totality of occupied units)에 대한 접근 가능한 자료를 일부 사용하여 품질의 변화정도에 대한 판단에 이를 수 있다는 견해를 지지하고 있다.

1976년에 비해 1993년 완성된 새로운 단신가구주택 속성들이 변모한 것은 품질수준 변화를 추가로 시사한다: 면적(제곱 feet)의 중앙값은

30% 증가하였고, 화장실은 2.0개에서 2.4개로 증가하였고, 중앙 에어컨 디셔닝을 갖춘 비율은 49%에서 78%로 증가하였으며, 화로를 하나 이상 갖춘 비율은 45%에서 63%, 그리고 차고를 갖춘 비율은 72%에서 84%로 증가했다.

우리는 이미 1976년과 1993년 사이 미국에서 실제 지불된 임대료평균이 CPI 임대료상승에 비해 1.0% 빠르게 상승했음을 판단한 바 있다. CPI에 바이어스가 없는 것으로 결론짓기 위해서는 평균적 임대단위의 품질이 그 기간에 걸쳐 매년 1.0% 증가, 또는 전체기간에 대해 18% 증가한 것으로 판단해야 한다. 검토된 증거에 따라, 우리는 평균적인 아파트규모 상승에 대한 추정치 하한을 20%로 본다. 이는 평방feet당 평균 임대료가 CPI에 비해 빠른 상승을 보이지 않았다는 결론을 지지할 만하다. 그러나 우리는 평방feet당 평균적인 아파트 품질 또한 향상된 데에 대한 설득력 있는 증거가 있는 것으로 본다. 지난 20년간 중앙 에어컨디셔닝으로의 전환은 매우 빠른 속도로 진행되었다. 아파트의 평균적인 품질개선을 가져온 기타 편의시설도 추가되었다. 특히 수영장, 헬스클럽, 야외 무료주차장, 그리고 기후(왜냐하면 아파트군이 남부기후대로 이동하여 거주자, 특히 나이든 거주자에 대한 겨울의 영향을 줄였기 때문이다) 등에 여기에 해당한다.

1970년 이후의 경우 CPI가 아파트 공간 1 평방피트당 임대료를 정확히 측정할 가능성이 높은 것으로 본다. 그러나 주거임대에 대한 척도는 평방피트당 아파트의 품질향상을 무시함으로써 상향 바이어스를 지니고 있다. 1인당 소득증가가 주택품질 향상으로 넘어가는 것은 당연하다. 왜냐하면 주택규모 및 품질이 반드시 영의 소득탄력성을 가져야 할 이유는 없기 때문이다. Gordon(1990)이 기록한 전기제품들의 품질개선은 주거부문에 적용된다. 왜냐하면 현재 대부분의 아파트에서는

비교적 최신 냉장고, 가스스토브 또는 오븐/가스레인지 세트, 식기세척기, 그리고 쓰레기 청소기 등이 제공되고 있기 때문이다. 이러한 전기제품들의 임대료 대등치는 상당할 것임에 틀림없으며, 주로 1955-60 이래로 새로운 아파트와 오래된 아파트에 모두 포함된 바 있다. 보수적으로 추정해보면 모든 전기제품의 임대가치, 중앙 에어컨디셔닝, 화장실 배수관 개선, 그리고 기타 편의시설들을 포함하여 평방피트당 아파트 품질 개선은 지난 40년간 총10%, 또는 매년 0.25%에 이르렀다. 이에 따라 표2는 주거 구성요소에 대한 CPI의 상향 바이어스를 0.25%로 기록하고 있으며, 이 또한 과소평가일 가능성이 다분하다.

1973년 이전에 대해서는, CPI지표가 품질변화로 설명되는 것 이상으로 하향 바이어스를 보일 수 있다는 증거가 있다. 예를 들어, CES의 근로자가구 연간 임대료지출 평균은 1950년의 \$444에서부터 1973년 \$1803,1 즉 4.06배로 증가한 반면, 조정을 거치지 않은 CPI 임대지표의 이와 대등한 비율은 단지 1.93에 그치고 있다. 이는 CPI 연간성장률이 매년 -3.24퍼센트 감소하고 있다는 해석으로 귀결된다. 1918년부터 1950년에 대한 동일한 비교는 매년 -2.82%의 차이를 나타낸다. 연간품질변동에 대한 척도없이 우리는 바이어스의 규모를 판단할 수 없다. 그러나 CPI 임대료지수가 1973년 이전에 상당한 하향 바이어스를 내포하고 있었을 가능성은 위에서 확인된 “Nordhaus 사고의 실험 문제”를 설명하는 데에 기여할 수 있다. 이는 바로 100년 또는 그 이상의 기간에 대해 현저한 CPI 바이어스를 거꾸로 추론할 경우 19세기 생활수준이 터무니없이 낮게 도출되는 것이다. 1918년, 욕조가 있는 아파트는 36%에 불과했고, 실내 번기가 있는 것은 61%에 불과했던 때에 비해서는 엄청난 품질 개선이 있었다는 사실이 명백하다는 점을 인지하면서도, 이 이슈에 대한 추가적인 판단은 1918년과 1973년 사이 아파트 품질변화에 대한 수량적 척도가 개발된 이후를 기약해야 한다. (Brown, 1994, Table 3.6A).

이제 주거지출의 여타 구성요소를 살펴보면, CPI가 연료 또는 전기의 가격을 부정확하게 측정한 것으로 간주할 이유는 없다. 이 상품들은 동질적이고, 가장 측정하기 쉬운 상품 또는 서비스에 속하기 때문이다. 그러나, 왜 사람들이 현대시대에 사는 것을 선호하고, 기꺼이 (대부분의 경우) 70년 전의 여건으로 되돌아가고자 하지 않는 이유를 생각하면, 가구 난방연료 성격의 변화가 그 이유를 산출하는 과정에 산입될 것이 확실하다. 1918년에, 미국주택 중 80%는 석탄과 나무로 난방이 되었는데, 이는 별도 저장 및 운반이 필요했고, 이를 통해 발생하는 불은 계속 지켜야 했으며, 난로(stove)는 청소가 필요했으며, 난로는 공기를 오염시키는 연기를 발생시켰다. 석탄과 나무로부터 다른 연료원천으로의 이행이 1970년대 초반에 대체로 완료되었기 때문에, 생활에 있어서 이 중대한 개선은 최근 CPI 바이어스 원천에는 포함시키지 않는다.

CPI의 주거에 대한 나머지 가중치는 각각 상대적으로 작은 비중을 차지하고 있는 무수히 많은 지출들에 적용된다. 여기에는 전화서비스, 쓰레기수거, 케이블 TV, 커튼, 가구, 침구류, 비디오 및 오디오 제품, 주요 가전제품, 그리고 기타 품목들을 다수 포함된다. “기타 재화”에 대한 CPI 가중치중 대부분은 시내 및 시외전화 서비스와 케이블 TV에 적용된다. 심지어 CPI가 각각의 품목가격을 올바르게 추적했다고 하더라도, 품질변화는 놓치게 될 것이다. 전화서비스 품질의 지속적인 개선(예를 들어, 공전(static)의 감소 및 음질의 투명성 향상), 편리성 증진(신용카드 공중전화, 항목별 청구서), 그리고 새로운 상품으로 간주되는 케이블 TV로 인해 달성된 월등한 화질향상 및 소비자 선택기회 증진 등이 이루어져왔다. 매월 프로그램 요금이 상당함에도 불구하고 미국 가구 중 60%를 넘는 비율이 케이블 TV에 가입해 있다는 사실은, 케이블 TV의 개발로 소비자잉여가 현저한 상품이 창조된 점을 시사한다. 보수적으로 추정할 때 우리는 이 범주와 관련된 품질 바이어스를 매 10년간 10%, 또는 매년 1.00%로 본다.

전기제품 및 라디오-TV 범주는 소비지출의 다른 어떠한 범주보다도 철저한 연구대상이 되어 왔다. 1947-83년 기간 전체에 걸쳐, 소비 보고서의 각 모델별 비교에 기초한 Gordon의 상세연구(1990, p.552)는 PCE deflator(이는 다시 CPI에 바탕을 두고 있다)에서 전기제품은 매년 3.22%, 그리고 라디오-TV는 5.94%의 상향 바이어스가 있음을 밝혀냈다. 이 중 1973-83 하위기간에는 해당비율이 차례로 2.83%와 4.69%이다. 이 비율들은 CPI에 있어서 표2에 의하면 오직 0.8%에 불과한, 소비중 지극히 작은 비율에 적용된다. 소비 전자제품만 해도, 즉 전기제품을 제외하고, 1994년에 \$55.9 billion의 연간 공장판매(즉 소매가격 인상 제외)를 기록했으며, 이는 총 개인소비지출의 1.25%에 달하는 금액이다. 1995년 최종 소비자 판매 중 PCE에서 차지하는 비중은 TV와 VCR을 포함한, 오디오 및 비디오 장비가 1.25%, 그 외에 전자제품이 0.55%, 그리고 PC 0.33% 등을 합하여 총 2.13%를 차지했으며, 이는 동일 제품이 CPI에서 차지하는 비중의 두 배를 거뜰히 넘었다.

개인 소비의 이 작은 단면은, 진료 행위를 제외해야 할 가능성이 있지만, 연간 바이어스비율의 가장 큰 원천이다. Gordon의 연구에 바탕을 둔, 전체 바이어스 추정치는, 품질변동 바이어스와 신상품 바이어스를 모두 포함하고 있다. 왜냐하면, 전체 바이어스에 대한 그의 추정치는 VCR이 1987년 CPI에 도입되기 이전인 1980년대 초에 매년 30-40%의 품질-조정가격의 하락을 경험한 사실을 반영하기 때문이다. 이와 유사하게, 최근 몇 년동안 소비자들이 구매하는 PC의 가격은 매년 적어도 25% 하락해왔는데, 이는 (CPI에 대해) 아무런 파장도 미치지 못한다. 왜냐하면 PC의 가구 구매가 CPI중 차지하는 비중은 기준시점 1982-84에 무시할 수 있는 수준이었기 때문이다.

이 부문의 전체 바이어스를 추정해보면 1년에 가전제품 3.0%, VCR 및 캠코더를 포함한 라디오-TV 4.0%, 그리고 PC는 15%등이다. 전기제품은 0.8%, 가전제품은 1.0%, PC는 0.4% 등 현행 경상가중치를 적용하면, 이 범주로 인해 총 CPI는 매년 0.10%의 연간 품질변화 및 신상품 바이어스를 더 갖게 된다. 표2에 기재된 숫자는 바이어스 추정치를 가중평균한 값이다. 그러나 총 CPI의 바이어스 숫자는 현행 경상지출에 대응하는 가중치를 바탕으로 하고 있으며, 표2에 기재되어 있는 가중치를 사용하지 않았다. 그리고 1994년 이전, 바이어스 숫자는 오직 전기제품 및 라디오-TV에 그 기초를 두고 있다. 왜냐하면 PC는 그 당시까지 주목받는 상품으로 부각되지 않았기 때문이다.

전기제품과 비디오 및 오디오 제품을 제외한 주거용 가구에 관해서는 지침을 제공할 만한 연구결과가 나와있지 않다. 가구, 직물 등이 광범위하게 제공됨에 따라 소비자들은 CPI에 포착되고 있지 않는 차원으로 상품, 직물 사이를 대체할 수 있게 되었다. 이 분야는 새로운 상품들이 여럿 존재해왔는데, 이전에 비해 얼룩 및 어린이 사고로 피해 입을 확률이 훨씬 적은 가구 및 직물이 포함된다. 이 범주는 또한 상당한 진전이 이루어진 비누와 세척제품들을 포함하고 있다.

**3. 의복.** 의복은 흔히 품질변화가 없는 것으로 간주된다. 그러나 청바지와 반바지, 고급 운동화류, 구김방지 합성직물, 그리고 경량이면서도 방수기능이 있는 우비 등을 포함하여, 소비자 후생을 높여주는 새로운 의복이 설새없이 등장하고 있다. 그럼에도 불구하고, 의복은 CPI가 하향 바이어스를 포함하고 있는 것으로 여겨지는 또 하나의 주요 분야이다. 하나의 하향바이어스 원천은 CPI 조사가격이 할인 판매되는 구형 모델의 가격하락을 추적한 후, (삭제 기법을 사용하여)신형 모델로 전환하면서 그에 따른 가격상승을 정확하게 기록하지 않았을 때 발생

하였다. 1980년대 중 후반에 수행된 개혁은 하향바이어스의 이 원천을 제거하였고, 의복 구성요소 내에서 일부 품질 조정을 위해 hedonic 가격기법으로 전환하였다.

CPI 의복지수는 우편주문 카탈로그 등의 자료원천으로부터 외부증거를 축적함으로써 상대적으로 쉽게 평가할 수 있다. 패션상품은 스타일 변화가 흔한 반면, 평균적 도시소비자들이 구입하는 실용복 제품은 우편주문 카탈로그를 통해 충분히 동일모델간 가격을 면밀하게 비교할 수 있을 정도로 품질변화의 빈도가 적다. 연간 또는 10년 주기 비교를 바탕으로 한 재차 점검을 결여한, 매월별 체계에 한정됨에 따라 CPI는 위에서 논의한 할인판매 상품들의 처리 등 측정오류에서 비롯된 끈질긴 압력에 노출되어 있다.

새로운 프로젝트에서 Gordon(1996)은 1965-93년에 걸쳐 동일 의복품목을 연도별로 수천 개 비교한 자료를 바탕으로, Sears 카탈로그로부터 의복가격지수를 축적하였다. 1985-93년 동안 Sears 의복지수 대비 CPI 비율은 연간 1.92%의 성장을 보였다. 1985년 이후 Sears 대비 CPI의복지수의 급속한 상승은 놀라운 사실인데, 이는 이 시기에 Sears는 Wal-Mart 및 기타 할인점들에게 시장점유를 빼앗기고 있었기 때문이다. 따라서 Sears 카탈로그 지수는 평균적인 미국 소비자가 직면한 진정한 의복가격의 상승을 과대 평가할 가능성이 있는 것으로 받아들일 만한 이유가 있다. 그럼에도 불구하고, 우리는 함축 바이어스를 Sears 지수에서 제시한 1.92%에서 더욱 삭감하여 1.0%로 상정하는 보수적인 접근방식을 채택할 것이다.

4. 교통. CPI의 교통 구성요소는 광범위한 이질적인 상품들로 구성되어 있다. 신규 자동차, 중고 자동차, 엔진 연료, 자동차 수리, 자동차 보험 및 등록, 그리고 대중교통, 주로 항공운임 등이 여기 포함된다.

교통 부문에서 다루어야 할 가장 중요한 질문들은 의무 안전장치(mandated safety)와 오염방지 장치 등의 가치, 그리고 신규 자동차에 대비한 중고자동차의 처리 등이다. Gordon(1990, p.364)이 1947-83기간에 대해 기록한 바와 같이, 신규 자동차의 실제 가격은 동일 품목의 CPI보다 훨씬 빠른 속도로 증가했으며, 1967년 이후 이 상대적인 증가는 신규 자동차에 대한 기존의 hedonic 회귀분석 등식으로 거의 설명되지 못했다. 1967년도 대비 1983년도 주요 가격비율은, 실제 가격이 289.9의 비율로 증가했으며, 자동차 CPI는 202.6비율로 상승한 것이다. 그리고 이 차이는 안전과 환경 품질에 대한 CPI 조정과, 연료비 절감에 대한 Gordon의 조정 등으로 설명이 되고도 남았다. 이로 인해 Gordon의 최종 자동차 지수 대비 CPI 상향 바이어스가 1967-1983에 이르기까지 매년 0.44 초래되었다.

그러나, Gordon은 CPI가 오염방지 장치를 가격인상보다는 품질개선으로 취급한 사실을 인정하였다. 우리는 의무적인 오염방지 장치들이 간접세와 유사한 것으로 확신한다. 휘발유세는 보다 나은 고속도로의 형태로 혜택을 제공할 수 있으나, CPI에서는 타당하게 조세인상을 생계비 인상으로 취급한다. 오염방지 장치들은 보다 깨끗한 공기의 형태로 이득을 제공하나, 조세의 경우에 유추하여 자동차 가격상승으로 취급되어야 한다. Gordon이 제시한 상세 정보를 이용하여 우리는 오염방지 장치에 대한 CPI 조절을 품질변화로부터 가격변동으로 전환하여 그의 지수에 대한 대안을 산출해낼 수 있다. 그리고 이는 1947-83기간동안 신규 자동차의 CPI가 매년 0.94%의 하향 바이어스를 지니고 있었다는

결론에 이르게 한다. 우리는 안전벨트와 충돌대비 범퍼 등 안전장치의 형태를 지닌 품질변화의 가치에 대해서는 이와 유사한 조절을 가하지 않는다. 왜냐하면 우리는 소비자들이 자신의 안전과 이 장치들 사이 관계를 오염방지 장치와 공기 품질간 관계에 비해 보다 직접적으로 볼 것으로 판단하기 때문이다. 따라서 우리는 자동차의 내구성 증가로 상쇄되는 0.94% 하향 바이어스를 정정할 것이다.

Gordon 연구의 마지막 해인 1983년 이래로 무슨 일이 일어났는가? Berry, Kortum, 그리고 Pakes(1996)는 1990년 연구 결론에서 자동차배기 규제기준이 1975년과 1979-81년 2단계에 거쳐 현저히 엄격해졌으나, 그 후에는 변화하지 않았음을 보여주고 있다. 이들은 또한 자동차의 마력과 가중치 변화 - 이는 1972-1983년 사이 67% 상승하고 1990년에 이르러 다시 13% 하락한다 - 에 대한 조절을 가하는 연료 효율성 지수를 개발한다. 연료 경제성에 대한 또 다른 연구(Sykuta, 1996)는 이 자료를 1994년까지 연장하고 “신규자동차 연료경제성은 1986년 정체기에 들어섰으며 그 이후 어떠한 방향으로도 1%이상 변화하지 않았다”(p.12)는 결론을 내리고 있다. 그는 연료가격이 낮아졌음에도 불구하고 연료경제성이 하락하지 않은 원인을 “보다 오래된, 보다 덜 효율적인 설계 및 기술들로 다시 돌아가는” 데에 따르는 높은 비용에 두고 있다. 따라서 오염방지 장비와 연료경제성 변화 모두 1980년대 중반 이후 신규자동차 CPI에 바이어스를 일으키는 중요한 잠재요인으로 작용했을 가능성은 없는 것으로 보인다.

그러나, Gordon의 연구와 CPI는 모두 전형적 신규자동차의 서비스수명 연장을 고려하지 않았다. 사용중인 자동차의 수명은 1970-83동안 연간 2.1% 증가하였으며, 이 기간은 일부 Gordon의 연구에서 다루어졌다. 그리고 1983-93 기간에 이는 1.3% 증가했다. CPI는 자동차도 주택을

취급하는 것의 연장선상에서, 귀속임대료를 기반으로 가격을 조사해야 한다. 만약 품질 향상, 유지보수에 대한 요구 감소, 녹(rusting)과 부식의 최소화를 가져다주는 기술변화를 통해 자동차의 유효수명이 연장된다면, 이 때 소비자들은 혜택을 누리게 된다. 1970-1993 동안 자동차 유효수명의 누적적 증가는 48%에 달하였다. 가격이 \$10,000이고 수명은 10년인 자동차를 살펴보자; 이는 연간비용이 \$1,000 소요되는 것과 대등하다. 이제 동일한 자동차의 수명이 48% 연장되어 14.8년으로 된 것으로 보자; 이에 따라 연간비용은 \$676로 감소했다.

자동차의 감가상각은(reduction in automobile depreciation) 자동차 소유의 사용자비용 중 일부에 지나지 않는다. 비용의 기타 수많은 요소들, 예를 들어 보험 및 수리비 등은 CPI에서 대범주 “기타 개인교통” 속에서 별도로 가격이 조사되는데, 여기에서 우리는 바이어스를 일체 자동배분하지 않는다. 사용자비용 중 감가상각 이외의 나머지 구성요소는 실질 이자비용이다. 현금, 할부금융, 그리고 리스 등을 포함하여 서로 대체가능한 자동차 대금지불수단을 적절히 조화시킬 경우, 우리는 실질 이자비용에 대한 타당성 있는 추정치는 10%인 것으로 생각한다. 자동차의 내구성 연장은 사용자비용을 매년 대략 20%(이는 10%의 실질이자비용과 10%의 감가상각률로 이루어짐)에서부터 16.7%(동일한 실질 이자비용과 6.7% 감가상각률의 합)로 감소시켰으므로, 1970-93 동안 사용자비용은 총체적으로 16.7% 감소했으며, 또는 매년 기하학적 비율로는 -0.79% 감소했다.

1983년 이후 자동차의 평균수명 증가율이 완화된에 따라, 우리는 이 효과를 배분하여 1970-83에는 매년 0.95% 비율로, 그리고 1983년 이후에는 0.59% 비율로 CPI 상향 바이어스에 기여한 것으로 본다. 1970-83 기간에 대해, 우리는 내구성 증가에 기인한 진정한 자동차 가격의 연

간 상향바이어스 0.95%를, 하향 바이어스에 대한 이전 추정치 연간 0.94%에 더한다. 이는 서로 정확히 상쇄되어 영의 순바이어스를 가져온다. 이어지는 1983년 이후 기간에 대해 우리는 CPI가 어느 방향으로도 편향을 보이는 것으로 간주할 만한 다른 이유가 없는 것으로 판단했다. 따라서 내구성 조절은 매년 0.59% 순 상향바이어스로 전환되었다.

이러한 산출과정을 평가하는 데에 있어서, 우리는 내구성 효과의 몇 가지 측면들이 순수하게 품질변화를 나타내는 것이 아니라, 오히려 자동차의 상대가격 상승으로 인해 사용자들이 자동차를 보다 오래 보유하거나, 또는 고속도로의 품질이 향상되는 등 다른 요인들을 반영할 수 있다는 사실을 인정한다. 이는 우리가 내구성 정정이 지나치게 크다는 사실을 암시할 수 있다. 그러나, 상쇄되는 오류가 보다 중요할 가능성이 있으며 이는 CPI 또는 우리의 정정에 반영되지 않은 품질향상의 중요한 원천을 나타낼 수도 있다 - 즉, 이는 J.D. Power 조사 및 여타 증거를 통해 측정된 바와 같이, 수입 및 국산 신규 자동차 모두에 있어서 결함 발생이 현저하게 감소했다는 것이다. 우리가 고려에 포함시키지 않은, 품질개선의 이 추가원천은, 우리가 추정한 신규 자동차 CPI의 바이어스가 보수적일 가능성이 높은 사실을 시사한다.

중고자동차의 CPI지수는 오래 전부터 상향 바이어스를 보이는 것으로 알려져 왔다. 이는 단순히 이 범주에 대해 품질 조정이 일체 적용되지 않았기 때문이다. 만약 신규 자동차와 중고차지수 사이 성장률 차이를 바이어스에 대한 하나의 척도로 받아들인다면, 1967-87 기간동안 상향 바이어스는 매년 2.44%에 이를 정도로, 매우 크다. 이를 영인 신규 자동차 상향바이어스에 더할 경우, 우리는 1967-87 동안 중고차의 총 바이어스가 연간 2.44%에 이른다는 결론에 도달한다. 1987년 이후 기간에 대해서는 CPI 내구성 바이어스를 중고차에도 적용하고 있다.

교통범주의 다른 구성요소에 관하여, 우리는 엔진연료를 동질적이고 측정이 용이한 것으로 간주한다. 그러나, 연료 및 관련제품에 대한 수많은 개선은 공기의 청정성 향상 및 내구성 향상과 같은 자동차의 개선에 기여해왔다. 공기 청정성 향상이 그 자체로서 가치가 있음에도 불구하고(제Ⅶ부 참조), 우리는 보다 휘발유 청정성 향상 의무를 굳이 품질변화가 아닌, 간접세로 취급한다. 여기에서 자동차에 대한 개선들이 반영되나, 이는 부분적으로 연료, 윤활유 등이 개선된 결과로 이해되어야 한다. CPI는 풀서비스 백화점과 부가서비스가 일체 없는 할인점의 경우와 마찬가지로, 풀서비스 및 셀프서비스 연료를 두 개의 서로 다른 상품으로 취급한다. 풀서비스에서 셀프서비스 엔진연료로 시장점유가 옮겨간 것은, 우리가 여기에서는 품질변화 바이어스에 포함시키지 않는, 일종의 상점대체 바이어스에 대하여 증거를 제공한다. 셀프 서비스로의 이행이 대체로 완료된 지난 10년간, 우리는 매년 0.25%의 CPI 상향 바이어스가, 휘발유 펌프에 내장된 자동 신용카드 단말기를 통해 얻게 된 편리함 증진 및 시간절약의 기여를 무시한 데에 기인하는 것으로 본다.

CPI의 중요한 범주인 “기타 개인교통”지출은, 주로 자동차 유지 및 수리와 자동차 보험으로 이루어져 있다. 우리가 아는 한 이 분야에 관련된 물가지수 바이어스가 존재한다는 증거는 없으므로, 여기에는 영의 바이어스를 부여한다. 그러나, 항공여행이 지배적인 부분을 차지하고 있는 대중교통 범주는, 1982년 이전에 CPI가 할인운임을 반영하지 못함에 따라 현저한 상향 바이어스를 포함하고 있었다는 사실이 잘 기록되어 있다. 이 바이어스에 대한 추정치는 Baily-Gordon(1988, p.416)으로부터 차용한 후, 이를 대중교통 범주에서 항공운임이 차지하는 비중(대략 2/3)으로 곱한다.

5. 진료. 진료범주는 전기제품 및 라디오-TV 제품과 마찬가지로 또는 이보다 높은 비율의 현저한 품질변화 바이어스가 존재할 가능성이 있다. 그러나 CPI에서 차지하는 비중은 전기제품 및 라디오-TV 제품에 비해 훨씬 높다. 이 가중치는 그 자체로서 논란의 여지가 많은데, 이는 그 7.4% 비중이 오직 소비자들이 즉석에서 지불하는 비용에 기반을 두고 있으며 노령 의료보험과 저소득 의료보험을 위한 지출, 또는 고용자-부담 봉급공제 등을 포함하고 있지 않기 때문이다. CPI의 주요 용도 중 하나가 실질소득, 보수 및 임금지표들을 “실질 단위”로 전환하는 데에 있으므로, 현행 가중치는 잘못된 것이다. 치료 가격 또는 기술변화는 총소득 및 보수에 해당하는 것과 동일한 비중으로 CPI에 영향을 미쳐야 한다. 즉, 이들은 훨씬 작은 비율인 CPI의 7.3% 비중이 아니라, 국민소득 및 생산계정 버전의 소비에서 16%에 달하는 총 치료청구서를 반영해야 한다.

7.4%의 CPI 가중치는 의약품, 전문 의료서비스(즉, 의사 및 간호사), 그리고 병원 등 세 가지 1차 범주에 적용된다. 처방약품에 대해서는 최근에 상당한 연구가 수행되었다. 예를 들어, Berndt, Cockburn, and Griliches(1996)는 항 우울증 약품의 가격을 연구했다. 이들은 1993-96 동안 매년 4.6% 상승하는, 노동통계국 방법론에 바탕을 둔 지수와, 매년 1.1% 상승하는, 자신이 선호하는 방법론에 기반을 둔, 이를 대신하는 지수 사이에 상당한 차이가 존재함을 발견하였다. 이 대체적인 지수는 대체적인 (Divisia) 가중치 체계를 사용하고, 새로운 품목을 보다 즉각적으로 반영하며, 소유권보장이 되지 않는 의약품과 특허약을 전적으로 동등하게 취급한다. 1995년 7월, CPI는 소유권보장이 되지 않는 의약품 처리방식을 전환했다. 따라서 특허약과 대등한, 소유권보장이 되지 않는 의약품이 판매될 경우 발생하는 가격하락이 모두 전적으로 가격변동에 반영되어, 다른 상품으로 취급되어 지수로부터 “연결

이 끊어지지(linked out)"않게 되었다. Berndt et. al. 연구 및 Griliches and Cockburn(1994)이 수행한 관련 연구에 기초하여, 우리는 1995년 이전 CPI 처방의약품에 매년 3.0% 바이어스가 있었던 것으로 결론짓는다. 1995년 7월 이후 소유권 보장이 되지 않는 약품에 대한 CPI의 새로운 취급방식을 통해 이 연간 바이어스 비율을 3.0%에서 2.0%로 감소된다. 그리고 나머지 바이어스는 (이전에 언급한 연구에 기초하여) CPI에 특정 약품 도입이 지연되는 것과, 시장점유가 확대되고 따라서 소비자들에게 가치를 더해주는 것으로 보이는, 새로 등장한 약품에 양의 가치를 부여하지 못하는 것에 그 원인을 돌릴 수 있다.

CPI 치료 구성요소의 주요 비중은 진료서비스, 즉 진료비(physician fees)와 병원비용에 적용된다. CPI는 보건진료 “산출”의 가치를 “투입”에 대비해 평가해보기 위해 시도하고 있지 않다. 생산자물가지수는 변형된 접근방법을 취하고 있는데, 이 방법은 투입의 가격을 산출한 후 다른 가중치를 부여하며, 최근(1995-96) 의사 및 병원 범주에 있어서 모두 CPI에 비해 매년 대략 2.0% 낮은 증가율을 보였다. 따라서 정부 지수는 이미 CPI에 매년 적어도 2%의 상향 바이어스가 있다는 중대한 증거를 제공하고 있으나, 여기서는 진료업무 및 기술에 대한 여러 가지 변화들이 무시되었기 때문에, 결론적으로 이보다 높은 바이어스 비율을 암시하고 있다. Cutler et. al. (1996, 표1)은 심장마비 치료비용의 투입 물가지수와, 이와 대체관계에 있는, 일정한 예상수명 증가의 달성비용을 반영한 “산출”지수들을 대조한다. 저자들은 1983-94에 걸쳐 매년 3.3% 증가하며 현재 CPI에서 축적하는 것과 동일한 종류의 “투입”지수를, 매년 1.1% 감소하는, 신규 심장마비 치료에 기인하는 예상수명 연장에 대한 보수적인 가치평가를 반영한 최종산출 지수를 대조하여, 매년 4.4%의 순 바이어스를 산출해냈다. Shapiro와 Wilcox는 1969-1993 동안 백내장 수술을 조사했다. CPI 방법을 모방한 “전형

(prototypical)”지수는 9의 인수로 증가하는 데에 반해 그들이 선호하는 대체 지수는 3의 인수로 증가한다. 이는 연간 4.57%의 상향 바이어스 비율을 내포하고 있다. Cutler et. al.과 Shapiro-Wilcox 연구가 상당히 다른 진료절차에 관한 것임에도 불구하고 서로 밀접히 연관되어 있는 사실은 상당히 주목할 만하다. CPI가 PPI에 비해 치료(산출보다는 오히려 투입 접근방식을 통한)에 있어서 1995-96동안 매년 2% 상승한 점을 염두에 두면, 우리가 진료범주에 부여한 상향 바이어스 비율, 연간 3.0%는 보수적일 가능성이 높다. 다양한 신규 진단 및 시험장비, 일상적인 치료절차의 고통 감소, 입원일수 감축, 그리고 병원 입원실의 품질 등을 고려하면 추가적인 품질변화가 존재할 가능성이 높다.

Cutler 및 기타에 의해 수행된 이 새로운 연구는 우리가 진료의 경제학을 이해하는 데에 있어서 중대한 진보를 위한 가능성을 열어주고 있다. 이 범주에 CPI는 미래 연구투자의 상당한 부분을 투자해야 할 것이며, 우리는 CPI가 보건진료 투입에 대해 가격을 조사하는 것에서 치료결과에 대한 가격을 조사하는 시도로 전환할 것을 강력히 지지하는 바이다.

**6. 여가생활.** 여가생활 범주는 상품과 서비스로 대략 균등하게 나누어져 있다. 상품은 신문, 잡지, 스포츠 장비, 그리고 장난감 등으로 구성되어 있다. 신문과 잡지에는 바이어스가 없으나(전자뉴스 서비스가 일부 사람들에게 편리함과 시의성을 제공함에도 불구하고), 스포츠 장비와 장난감에는 Gordon이 전기제품에서 발견한 3.0% 보다 다소 낮은 비율, 즉 1년에 2.0% 의 바이어스가 존재하는 것으로 가정한다. 이는 Nintendo 오락기 등 전자장난감의 15% 와 - 이는 우리가 PC에 적용한 15% 비율에 가까운 것으로 볼 수 있다 - 동물인형 및 비전자식 보드 오락기 등을 포함한 기타 장난감의 바이어스 비율 영을 평균한 값을

의미한다. 여가 서비스는 클럽 회원권, 입장료와 레슨비 등으로 구성되어 있다. 스포츠 행사 관람에 있어서 돔 스타디움과 보다 편안한 의자를 갖추으로써 안락함이 증가한 데에는 의심의 여지가 없으나, 여가 서비스 가격에 대한 CPI 척도에는 바이어스를 부여하지 않았다.

**7. 기타 상품 및 서비스.** 담배, 세면용품 및 화장품 등의 상품구매가 상점대체 바이어스 영향을 받았을 가능성은 있는 반면, 이 때 품질변화를 간과했거나 신상품 바이어스의 비중이 높았던 것으로 생각할 만한 이유는 없다. 그러나 이 범주는 헤어드라이어 등 소형 개인장비를 작은 비율로 포함하고 있으며, 이 품목들이 대형장비와 동일한 규모의 바이어스, 즉 1년에 3.0% 하에 있는 것으로 간주하는 것이 합당하다. “기타”범주의 나머지 비중의 상당부분은 교과서와 주로 대학 등록금에 해당하는 수업료로 구성되어 있다. 그러나, 이 범주의 대략 1/10은 개인 금융서비스로 이루어져 있는데, 이는 급속도로 진행된 기술변동, 특히 ATM기계와 all-in-one 현금관리 계정의 광범위한 확산으로부터 영향을 받아왔다. 보수적인 추정치 2.0%를 개인금융서비스에 채택하고 이를 전체 범주의 1/10에 적용할 경우 연간 바이어스 비율 0.2%에 도달한다.

### 품질변동에 대한 결론

표2에서 가중치와 바이어스 비율을 차용한 결과 품질변동 바이어스의 최종 추정치는 대략 1년에 0.6% 합계로 도출된다. 이는 중간보고서에서 추정된 품질변화 및 신상품 바이어스 통합범주의 바이어스비율인 1년에 0.5%와 Shapiro 및 Wilcox(1996c)의 0.358%에 비해 약간 높다. 우리의 추정치가 보다 높은 것은 기존문헌을 훨씬 심층적으로 연구했으며 표2에 포함된 별도의 27개 범주 각각에 대해 영향을 미치는 요소를 고려한 데에 기인한다.

표 2 1995년 12월, CPI-U, U.S. 도시 평균의 구성요소간 상대적 중요성, 그리고 선정된 기간에 대한 품질변동 및 신상품 바이어스 추정치

주요 및 선정된 비 주요 구성요소	상대적 중요성(%)	선정된 기간에 대한 품질변동 추정치(연간비율)		
1. 음식 및 음료	17.332			
농산물 이외 집에서 소비하는 음식	8.543	0.30 (1967-96)		
신선한 과일 및 채소	1.337	0.60 (1967-96)		
외식	5.886	0.30 (1967-96)		
알콜 음료	1.566	0.15 (1967-96)		
2. 주택	41.346			
주거	28.289	0.25 (1976-96)		
연료	3.792	0.00		
기타 재화, 전화 포함	3.222	1.00		
전기제품, 전자제품 포함	0.806	3.6 <sup>a</sup>	5.6 <sup>b</sup> (1994-96)	
기타 가구	2.693	0.30		
가사장비	1.116	0.00		
가사서비스	1.482	0.00		
3. 의복 및 유지	5.516	-0.95 (1965-85)	1.00 (1985-96)	
4. 운수	16.953			
신규 자동차	5.027	0.00 (1970-83)	0.59 (1983-96)	
중고 자동차	1.342	2.44 (1967-87)	1.59 (1987-96)	
엔진 연료	2.908	0.00 (1974-84)	0.25 (1984-96)	
기타 개인교통	6.153	0.00		
대중교통	1.523	2.66 (1972-77)	4.60 (1977-82)	0.00 (1982-96)
5. 진료	7.362			
처방 의약품	0.891	3.00 (1970-95)	2.00 (1995-96)	
비처방 의약품 및 의료장비	0.391	1.00		
전문 진료서비스	3.465	3.00		
병원 및 관련서비스	2.257	3.00		

의료보험	0.358	0.00		
6. 여가생활	4.367			
상품	1.975	2.00		
서비스	2.392	0.00		
7. 기타 상품 및 서비스	7.123			
담배, 담배 제품	1.610	0.00		
개인용 장비	1.170	0.90		
개인용 및 교육비용	4.342	0.20		
합 계	100.000	0.612(1996년)		

품질변동 바이어스가 증가했는가 아니면 감소했는가? 표2는 이 이슈에 대한 몇 가지 직관을 제공한다. 1980년의 구체적인 예를 보면, 표2에 기재된 바이어스 비율의 합계를 통해 품질변화 및 신상품에 기인한 CPI 상향 바이어스가 1년에 0.488%임을 알 수 있으며, 이는 1996년의 0.613%와 비교된다. 격차는 주로 이전 의복의 하향바이어스가 역전되고 PC의 역할증대에 기인한 전기제품/라디오/TV 구성요소의 상향바이어스 증가로부터 파생되었다. 중고자동차, 항공운임, 그리고 처방 의약품의 상향바이어스 완화는 상향 바이어스 심화의 이 원천들을 부분적으로 상쇄하고 있다.

## VI. 각 종류별 및 총 바이어스 추정치

CPI는 생계비지수가 아닌, 고정 가중치 지수이며, 수정 라스파이레스 산식을 통해 시행된다. 이는 CPI의 목표와 수많은 예정된 용도들 사이 즉각적인 충돌을 야기한다. 예를 들어 CPI는 민간 계약, 조세구간 그리고 사회보장 등 정부 이전지출 프로그램을 연동시켜 수혜자에게 생계비 변동분을 보상하기 위해 이용된다. 고정가중치 지수는 물가변동이 생계비에 미치는 영향을 과장하며, 이는 소비자들이 상품대체를 통해 충격전체를 받는 일은 회피할 수 있다는 사실을 반영하지 못하기 때문이다.

CPI 바이어스의 평가를 위해서는 고려의 초점을 생계비지수에 맞출 필요가 있다. 생계비지수의 근사치는 CPI에 있는 여러 종류의 바이어스에 관한 연구결과를 종합함으로써 도출된다. 이 장의 목적은 이 연구를 요약하고 이 바이어스들이 CPI 내에서 갖는 중요성을 평가하는 것이다. 우리의 추정치들은 주로 본 위원회가 구성되기 이전, 상당 부분 노동통계국에서 수행한 연구에 의존하고 있다.

1995년 9월 15일자 중간보고서는, CPI의 바이어스에 대한 애초의 추정치를 제시했다. 우리는 최근 몇 년간 전체 바이어스가 1년에 1.5%에 해당했으나, 그 당시 노동통계국에서 예정하고 있던 CPI 방법론 변경을 통해 여기서 1년에 0.5% 규모의 바이어스가 제거되고, 따라서 바이어스는 1년에 1.0% 포인트를 향하여 감소할 것으로 추정하였다. 이제 우리는 노동통계국에서 1996년 3월 29일 발표한 CPI 변경과, 신상품의 등장 및 기존 상품의 품질변동이 미치는 파장에 대한 새로운 추정치를 반영하도록 추정치를 개정하였다. 노동통계국은 매년 총 0.24% 포인트에 달하는 이 바이어스를 일부 제거했으며, 이로 인해 우리의 바이어스 추정치는 0.25% 포인트 상승했다. 그 외에도, 우리는 신상품/품질변동 바이어스 추정치를 1년에 0.10% 포인트 높이 개정하였다.

CPI 바이어스를 평가하는 데에 있어서 2 종류의 대체 바이어스를 서로 분리하는 것이 결정적이다. 첫째, 노동통계국은 수정 라스파이레스 산식에 바탕을 둔 고정가중치 지수를 이용하여 44개 지역 207개 품목의 물가지수들을 전국 CPI로 통합한다. 가중치는 소비지출조사(CES)로부터 추론되며 개별 가구조사를 반영한다. 우리는 이 수준의 대체바이어스를 상급 대체바이어스(Upper Level Substitution Bias)로 지칭한다. 이 바이어스는 노동통계국에서 사용하는 수정 라스파이레스 산식과, 대체바이어스가 없는(대략 없다고 할 수 있는) Tornqvist 지수간 격차

로 측정된다. 대부분의 추정치는, 노동통계국의 술한 추정치를 포함해서, 0.2와 0.25 주위에 군집해 있다. 위원회에서 구할 수 있는 최근 추정치는 이제 갓 제출된 미공표 정정사항들을 노동통계국의 이전 연구에 반영했으며, 1988-1995기간에 대해 대략 1년에 0.15%포인트 평균 바이어스를 보인다. 이 신규 결과를 분석할 시간은 갖지 못했으나, 보수적인 입장을 취하기 위해, 우리는 이 수치를 채택한다.

대체바이어스의 두 번째 종류는 대략 71,000개 상품 및 서비스의 가격 관측치와 CPI의 주거 구성요소 가격을 207개 품목과 44개 지역의 지수로 통합하는 데에 기인한다. 우리는 이 수준의 대체바이어스를 하급 대체바이어스(Lower Level Substitution Bias)로 지칭한다. 수집예정인 물가는 확률표본을 통해 선정된다. 확률은 CES와 소매점포 POPS로부터 추론되며 기준시점의 소비지출 중 품목 및 지역이 차지하는 비중을 반영하는 데에 그 의도가 있다.

1978년부터 노동통계국은 각 품목 및 지역범주내의 물가지수 추정에 표본교체 절차를 도입하였다. 동시에 가격들은 수정 라스파이레스 산식을 반영한 방법으로 종합되었다. 이 두 가지 변경은 지수에 하위 수준 대체바이어스가 상당수준 포함되는 결과를 가져왔으며, 이 사실은 15년의 지연을 거친 후 Reinsdorf(1993)에 의해 발견되었다. 노동통계국에서는 이 이슈를 심층적으로 연구했으며 1995년 1월을 기점으로 이를 처리하기 위한 조치를 취한 바 있다.

1992년 6월부터 1994년 12월에 이르는 30개월 기간을 포괄하는 CPI자료에 대해, Moulton과 Smedley(1995)는 노동통계국이 하위 수준의 통합 과정에 사용한 수정 라스파이레스 산식과, CPI의 비주거 구성요소들의 가중 기하평균 산식 사이 격차가 1년에 0.49% 포인트인 것으로 추정

하였다. 이 차이는 라스파이레스 산식의 바이어스에 대한 추정치이며, 이는 Shapiro와 Wilcox(1996)가 기하평균이 내재된 생계비 지수의 불편 추정치임을 증명했기 때문이다. Armknecht, Moulton, 그리고 Stewart(1995)는 자가소유의 귀속임대료를 1년에 0.50% 포인트로 추정했다. 현재 노동통계국은 1990년 물가자료를 시작으로 이 방법론에 기초한 지수작성이 실현 가능한지 여부를 실험하고 있다. 하급의 통합단계에서 수정 라스파이레스 산식을 가중 기하평균 산식으로 대체하는 것은 생계비지수를 향한 중요한 단계이다.

1995년 1월 노동통계국은 가정내 소비음식, 자가소유의 임대료 대등치, 그리고 처방의약품의 물가변동을 자동배분(imputation)하는 방법을 개선했다. 이 변경은 Armknecht, Moulton, 그리고 Stewart(1995)에 의해 소개되고 있다. 가정내 소비음식 품목들에 대해서는 “평활화(seasoning)”로 지칭되는 절차도입이 포함되었다. 평활화 기간은 지수에 도입하기 이전, 각 신규표본 관측치의 비중을 측정하기 위해 필요한 자료수집 기간이다. 이 기간은 가정내 소비음식 품목에 있어서 3개월로 연장되었으며, 이를 통해 이 품목들의 가중치와 궁극적으로 CPI에 사용된 가격 사이 연결고리가 끊기게 되었다.

1996년 3월 29일, 노동통계국은 1996년 6월 CPI에서 효력을 발생한다는 전제로, 평활화 절차가 모든 비주거품목에 확대되는 것으로 공표했다. 이 공표는 거주임대료와 자가소유 임대료 대등치가 1995년 1월 도입된 변경 결과, 더 이상 표본교체 절차와 관련된 바이어스 하에 놓여 있지 않음을 지적했다. 이와 더불어, 노동통계국은 (드문 예외를 제외하고는)품목의 수명 전체를 통틀어서 교체품목의 가중치를 일정하게 유지하는 것으로 선언했다. 노동통계국은 1995년 1월 변경에 기인한 바이어스 감소를 1년에 0.14%포인트로 추정한 반면, 비주거품목에 대

한 1996년 변경은 추가로 1년에 0.10%포인트 바이어스 감소를 가져올 것으로 예상했다. 이로 인해 나머지 하급 대체바이어스의 Moulton-Smedley 추정치는 1년에 0.25% 포인트로 감소한다. 1996년 7월 16일, 노동통계국은 CPI의 병원 및 관련서비스 구성요소 분류 및 정의에 대한 변경을 도입했다. 이는 CPI의 이와 같은 중요요소 측정을 개선할 의도로 도입되었으나, 바이어스 추정치 감소를 동반하지 않았다. 따라서, 우리는 하급 대체바이어스에 대한 점추정치를 1년에 0.25% 포인트로 한다.

노동통계국의 목표는 일정한 품질의 상품 및 서비스를 측정하는 것이다; 그러나, 상품들이 시장 공간으로부터 사라짐에 따라 이를 여타 상품들로 대체하는 일은 불가피하다. Armknect와 Weyback(1989)는 노동통계국이 품질변동에 대한 조정을 행하기 위해 사용한 방법들을 요약했다. 신제품 바이어스 및 품질변동 바이어스의 제거는 생계비 측정에 있어서 필수 불가결한 단계이다. Hausman(1996)의 아침식사용 시리얼에 대한 연구, Griliches와 Cockburn(1994)의 처방의약품에 대한 연구, Pakes, Barry, 그리고 Levinsohn(1993)의 신규 자동차에 관한 연구, 그리고 제V부에서 설명한 다수의 기타 연구들을 통해 신상품 바이어스에 관한 중요한 실증연구가 이루어졌다. 우리는 신상품 및 기존 상품의 품질변동에 기인한 총 바이어스를 1년에 0.6% 포인트로 추정하였다.

Reinsdorf(1993)는 신규상점 대체바이어스에 대한 주된 경험적 증거를 제시하였다. 이것은 CPI가 포괄하고 있는 표본에 교체되어 들어가는 상점들과 교체되어 나가는 상점들 사이의 일정한 음식 및 연료품목의 가격을 비교한 데에 바탕을 두고 있다. 그는 바이어스를 1년간 25% 포인트로 추정했다. Lebow, Roberts, 그리고 Stockton(1994)은 상점대체

바이어스의 영향을 받을 지수의 구성요소들을 확인함으로써 CPI 전체에 대한 이 추정치를 추론했다. 이 구성요소들은 CPI의 40%에 달하므로, 상점대체바이어스는 1년간 0.1% 포인트이다.

이 별도의 바이어스들은 대체로 누적적이며, 진정한 인플레이션율의 완곡한 진동으로부터 독립되어 있을 가능성이 높다. 따라서, 1% 포인트 바이어스는 만약 CPI 변동을 통해 인플레이션이 3%로부터 5%로 상승할 경우, 실제 인플레이션은 2%에서 4%로 상승했을 가능성이 높은 것을 의미한다. 바이어스는 주로 인플레이션율의 변동이 아닌, 그 절대치에 영향을 미치는 사실을 주목하라. 매우 높은 수준의 인플레이션에서는 바이어스가 증가할 가능성이 있으나(상점 및 상품 대체가 증가할 것으로 가정할 수 있다), 현재로서는 이 이슈에 관련된 증거를 확보하고 있지 못하다.

표3은 CPI 바이어스에 대한 우리의 평가를 요약하고 있다. 이는 가능한 가장 좋은 증거에 기초를 둔 점추정치 뿐만 아니라 전체 바이어스의 현실성 있는 구간을 포함하고 있다. CPI 전체 바이어스에 대한 우리의 추정치 평균은 1년에 1.1% 포인트이며 그 구간은 0.8에서 1.6% 포인트이다.

노동통계국은 1998년 1월, CPI가 1993-1995 소비지출조사의 새로운 지출가중치를 포함하게 될 때, 이정표가 될 만한 개정을 준비하고 있다. 그러나 노동통계국이 수정 라스파이레스 산식은 유지함에 따라, 우리의 바이어스 추정치들은 개정된 CPI에 그대로 넘어갈 것이다. 그뿐 아니라, 노동통계국은 끊임없이 측정 개선과 이미 밝혀진 결함들의 치유를 위해 CPI에 중대한 정정을 도입해왔다. 그러나, 이들 개정이 이정표가 될 다가오는 개정과 함께, 수정 라스파이레스 체계를 채택함으로써 인해 CPI와 생계비지수 사이에는 중요한 차이점들이 존속할 것이다.

표3 : CPI 기초 생계비 척도의 바이어스에 대한 추정 (1년당 퍼센트 포인트)	
바이어스의 원천	추정치
상급 대체	0.15
하급 대체	0.25
신상품/품질변동	0.60
신 상품	0.10
총 계	1.10
현실성 있는 구간	(0.80-1.60)

CPI의 상급 대체바이어스는 다가오는 CPI의 이정표 개정 이후에도 존속할 것이다. 왜냐하면 노동통계국은 수정 라스파이레스 산식의 유지를 계획하고 있기 때문이다. 둘째, 노동통계국은 1978년 기초수준 합계에서 표본교체절차와 수정 라스파이레스 지수를 도입함에 따라 상당부분 초래된 하급 대체바이어스의 일부인, 소위 산식 바이어스(formula bias)를 감소시켰다. 그러나, 1995년 1월과 1996년 3월 공표된 변경조치의 목적은 수정 라스파이레스 산식 적용을 개선하는 것으로, Moulton과 Smedley(1995)가 수치화한 하급 대체바이어스를 제거하는 것은 아니었다. 마지막으로, 신상품, 신상품, 그리고 품질변동 바이어스는 노동통계국이 이미 공표한 변경 또는 이정표가 될 개정으로부터 영향을 받지 않는다.

요약을 위해 우리는 노동통계국의 CPI 개정조치들과 바이어스 규모에 대한 새로운 발견들이 축적된 것을 반영하도록 중간보고서에 제시된 추정치를 개정하였다. 우리의 주된 결론은 수정 라스파이레스 산식에 의해 부과된 제약들로 인해 CPI가 생계비 측정에 부적합하게 되었다는 것이다. 상급 수준 합계에서 Tornqvist 산식과 하급수준에서 가중 기하 산식을 결합함으로써, 노동통계국은 두 종류의 대체바이어스를 모두 제거할 수 있다. 그러나 이 변경들만으로 신제품, 신상품, 품질변화 바이어스 조정을 모두 포괄할 수는 없다. 이 바이어스들에 대한 조정은 생계비측정을 위해 결정적인 중요성을 지닌다.

표2는 2006년에 이르기까지 CPI-U에 관한 CBO 예측치에 매년 1.1% 포인트 바이어스가 미치는 복합적인 효과를 묘사하고 있다. 특정 연도의 1.1% 포인트가 작은 규모로 보일 수 있는 데에 비해, 누적적으로 매년 연속되면 합해져서 12년 동안 14%에 해당하는 상당한 규모에 이르게 된다.

## Ⅶ. 기타 이슈

### A. 노령인구를 위한 별도의 지수?

실제는 아닐지라도, 원칙적으로는 모든 가구 각각에 대해 실제 소비바구니와 지불가격을 바탕으로 별도의 생계비 지수를 개발할 수 있을 것이다. 위에서 지적한대로, 종합 지수들은 대표적 소비자를 나타낸 자료를 사용한다. 일부에서는 인구 내의 서로 다른 집단들이 CPI 변화로 기록되는 것에 비해 보다 빠르거나 보다 느린 생계비 상승을 경험할 가능성이 높은 것으로 주장해왔다. 오늘날에 이르기까지 우리는 이에 대한 강력한 증거를 찾지 못했으며, 오히려 2개의 연구는 인구 그룹별로, 예를 들어 지역 또는 나이별로 지수를 분리하는 것이 측정된 생계비 변화에 거의 영향을 미치지 못함을 시사한다. 이에 대한 추가 작업은 향후 과제로 남아있다. 특히, 지출 중 비중이 아닌, 실제 지불된 가격은 차이가 날 수 있다.

서로 다른 소비바구니를 초월하여, 바이어스 원천에 관한 우리의 분석이 대표적인 또는 평균적인 소비자에게 해당된다는 사실을 이해하는 것이 중요하다. 일부 소비자들은 여타 소비자들에 비해 보다 많은 상품을 대체할 것이며, 일부 사람들은 대체바이어스가 보다 클 수 있고,

다른 사람들은 보다 작을 수 있다. 아마도 가장 중요한 것은, 품질변동과 신상품 도입으로 인한 혜택이 인구 전체에 대해 불균등하게 확산될 수도 있다는 사실이다. 이와 유사하게, 일부 사람들은 할인점으로부터 보다 많은 이득을 취하는 경향이 있을 것이며, 다른 사람들은 이 경향이 덜할 것이다. 어떤 이들은 예를 들어 핸드폰의 혜택을 발빠르게 누릴 것이며, 반면 다른 이들은 수년 또는 수십년간 기다리거나, 아예 사용하지 않을 수도 있다. 이는 우리가 바이어스의 이와 같은 특정 원천에 관한 점추정치에 대해 매우 조심스러운 입장을 취해온 또 하나의 이유에 불과하다.

## B. 삶의 질에 대한 보다 광범위한 고려

모든 변화가 긍정적인 것은 아니고 모든 변화가 모든 사람들에게 긍정적인 것도 아니다. 품질 변화의 중요성을 주장하는 과정에서, 우리는 손익(ledger)의 반대편도 살펴보아야 한다. 적어도 3가지 종류의 변화를 고려해야 한다: 1. 신상품들은 인구의 어떤 하위집단에서 아직 소중하게 여기는 구제품들을 축출할 수 있고 또는 이와 동일한 것으로, 규모의 경제 상실로 인해 이들 구제품이 현저한 가격상승을 경험할 수 있다. 이는 측정이 되는 한, 물가지수 구성에 있어서 새로운 문 제점으로 나타나지는 않는다. 2. 기존의 상품 또는 서비스 품질이 무너질 수 있다. 이는 흔히 속설로 “물건을 예전처럼 만들지 못해”라는 불평이 존재함에도 불구하고, 덜 흔한 현상이다. 가장 두드러지는 예는 아마도 의료보험의 구입을 통해 계약으로 예정된 서비스들이 현저히 악화된 것으로 받아들일 가능성이 높은, HMO가 도입한 진료접근에 대한 규칙 강화일 것이다. 그러나 이것이 최근 몇 년간의 의학 진보를 상쇄했을 가능성은 거의 없다. 심장동맥수술(bypass operations), 위궤양 치료, 또는 백내장수술 등의 진전을 포기하면서, 오늘날의 치

료 접근권 제한을 예년의 기술에 대한 보다 자유로운 접근으로 맞추고자 하는 사람은 거의 없을 것이다. 가장 큰 영향은 우리의 물리적, 사회적 및 경제적 환경변화로 인하여 이전에 달성했던 효용수준을 쫓아가기 위해 점점 더 높은 지출이 부과된다는 사실이 가장 큰 파장을 가져올 수 있다.

그러나, 보다 추운 겨울, AIDS 출현, 또는 범죄율 상승 등의 사건들이 물가지수의 정의에 포함되어야 하는지는 분명하지 않다. 예상하지 못한 날씨 변화에 기인한 지출변화는 에너지 가격이 상승하는 한도 내에서 물가지수를 상승시켜야 하며, 소비된 수량이 물가지수에 영향을 미쳐서는 안 된다. 만약 사건이 지속된다면 후자는 궁극적으로 지수내의 상품가중치를 변경시키게 되나, 이는 또 다른 사항이다. 만약 우리가 보건물가지수를 기준시점의 보건수준 달성에 필요한 지출수준으로 정의한다면, AIDS의 증가는 보건물가지수를 상승시킬 것이다. 그러나 이 구성요소가 “생계비”의 실질 상승을 나타냄에도 불구하고 지수 산식의 구성요소로는 적합하지 않을 수 있다. 왜냐하면 이러한 물가상승을 위해 퇴직자들에게 “보상”을 해야 하는 젊은이들 사이에는 “이득”이 없기 때문이다.

AIDS의 출현은 실제로 생계비 상승을 가져왔다. 만약 최초 환경악화가 애초 측정체계에 반영되지 않았다면 뒤이은 AIDS 치료의 진전을 긍정적인 품질변화로 반영하는 것이 부적절할 것이다. 이와 유사하게, 자물쇠 및 기타 안전장치의 품질개선을 반영할 경우 이와 같은 방어적 소비투자들로부터 얻은 “이득”을 과대 평가할 수 있다.

이 변화들을 완전하고도 정확하게 설명하는 것이 불가능한 한편, 위에서 인용된 발전에 내포된 ‘악화요인(bads)’들이 대두됨에 따라 여기 포

함된 ‘개선요인(good)’이 모두 상쇄될 가능성은 희박하다. 주요 관심 및 측정불량 대상인 환경, 보건 및 범죄에서, 앞의 두 분야는 발전을 거듭해왔으며, 최근 몇 년간, 마지막 분야는 안정적 수준을 유지하는 데에 근접해왔다. 따라서 측정노력을 소비 “바구니”를 구성하는 것이 무엇인지에 대한 현재 개념을 초월하는 범위로 확장하도록 권장하는 한편, 이미 조사한 분야들(현재 CPI 내용으로 정의된, 소비의 보다 전통적인 구성요소)에 있어서 측정이슈에 대한 우리의 결론들을 약화시킬 강력한 이유는 없는 것으로 본다.

산업혁명은 광범위한 공기 및 수질오염을 야기했으며, 1950년대와 1960년대에 이르기까지 이는 과연 부정적인 요소였다. 그러나 그 후 가구난방의 지배적인 에너지원천이 석탄으로부터 천연가스로 전환되고, 연료와 자동차의 청정성이 개선되며, 환경규제가 부과됨에 따라 공기 및 수질에 대한 여러 종류의 오염원이 현저하게 감소했다. 난방 연료의 전환은 또한 석탄통이 저소음 자동 천연가스 전송으로 대체되면서 집안주위의 청결과 편리함을 대폭 증진시켰다. CPI는 암묵적으로 자동차의 의무적인 오염방지 장치들로 인해 가능하게 된 공기의 품질향상을 높이 평가한다. 왜냐하면 이는 의무 오염방지장치를 가격상승이 아닌 품질향상으로 취급하기 때문이다. 우리는 CPI가 추가 오염방지 장치로 인한 가격변동을 품질향상이 아닌 가격상승으로 취급하도록 권장하였다. 그러나 본 보고서에서 품질변화 바이어스에 대한 우리의 추정치에 포함된 이 개념에는 공기 품질의 개선이 측정되지 않은 채로 남아있다. 이는 CPI 바이어스의 최종 추정치가 과소 평가되는 원인이다. 더욱이, CPI는 일관성이 없으며, 이는 전기, 철강, 종이 및 기타 제품들의 비용상승 중 일부 또한 환경규제에 기인하고 있기 때문이며, 자동차 이외 제품에 대한 규제로 가능하게 된 공기 및 수질 향상의 혜택을 여기 반영하지 않았기 때문이다.

범죄는 또 다른 종류의 외부성이다. 범죄피해율은 과거 1970년 인구당 0.096개 사건으로부터 1981년 0.098개 사건으로 증가해왔다. 그 이후에는 그러나 피해율이 1992년(구할 수 있는 가장 최근 수치)에 0.074개 사건으로 하락했다. 폭력범죄의 비율은 1970년의 0.025로부터 1981년의 0.028로 상승했으며, 그리고 1992년에는 0.026으로 약간 하락했다. 1992년 이후 여러 개의 주요 미국 도시에서 살인율이 급격한 하락을 보인 데 대해 광범위한 주의 관심이 집중된 바 있다.

부정적인 요소들을 추가로 찾아보면 이혼, 사생아, 그리고 핵가족의 역할감소 등 사회적 이슈들이 아마도 가장 중요할 것이다. 이혼율은 1970년과 1980년 사이 50% 증가했으나 그 후 안정적인 수준을 유지하고 있다. 1980년과 1992년 사이 자살율은 안정적이었으나 연령-조정 사망률은 14% 하락했으며, 유아사망률은 1970년과 1992년 사이 58% 하락했다. 탄생시점 예상수명이 1970년 70.8세에서부터 1993년 75.5세로 상승했다는 사실이 아마도 가장 중요할 것이다. 주된 부정적 측면은 미혼모에게서 탄생한 아이 비율이 18.4%에서 30.1%로 증가한 사실이다. 그 외에 직장의 불확실성 증가에 대한 인식, 불평등 증가의 가능성, 그리고 단지 고등학교 교육만을 받은 근로자들을 위한 취업기회 감소 등 무형의 부정적 측면들이 있을 수 있다.

긍정적인 측면에서, 품질변화에 관한 우리의 논의가 전적으로 포괄할 수 없는 방식으로 제품개선이 이루어진 사실은 의심의 여지가 없다. Gordon (1990, p.38)은 자신의 책도가 포착하지 못한 품질의 몇 가지 차원을 지적하고 있는데, 이 중 몇 가지는 제트 비행기의 진동감소 및 속도 향상, 전기제품 및 자동차의 신뢰성 향상, 가구 및 자동차 오디오 장비의 음질 향상, 가구용 연장(power tools) 및 강력 잔디깎이의 안전 장치 향상, 실내 에어컨디셔너의 소음, 무게, 그리고 설치비용 감소, 그리고 “컬러TV의 측정 불가능한 정도의 화질 개선” 등이다.

1870년대 마력을 동력으로 하는(horse-drawn) 시골 경제에서부터 오늘날의 현재 경제에 이르기까지 미국 생활수준 변화의 여러 측면들은 수십년 전에 이미 일어났으며 현재 CPI 바이어스의 원천은 아니다. 속도 및 유연성과 더불어, 동물 배설물 제거 등이 자동차가 주로 기여한 부분으로 평가되나, 이는 이미 제2차 세계대전 이전에 달성되었다. 전기제품을 통해 가사노동의 고역이 감소된 것도 1960년대에 대체로 달성되었다. 그러나 일부 진전은 지속적으로 이루어져왔다. 피아노(player piano)로부터 라디오로, 흑백TV에서부터 이후의 컬러TV, VCR, 그리고 현재는 케이블 TV, 위성TV로, 그리고 World Wide Web과 PC에서 이용 가능한 기타 기능들에 이르기까지 꾸준한 이행이 이루어져왔다. 또한, 중앙 에어컨디셔닝의 급속한 확산으로 인해 상당수의 가구들이 남부 그리고 남서부 주(州)로 이동하게 되었다. 수백만의 인구가 이 옵션을 선택한 바 있다. 이들은 보다 따뜻한 겨울과 인공으로 냉방된 여름을 즐기기 위해 “발로 투표(voting with their feet)”했다.

전체적으로 우리는 만약 악화될 경우 한 때 누렸던 삶의 질을 점차 손상시킬 가능성이 있는, 부정적인 것으로 추정되는 요인들(오염, 범죄, 자살, 이혼)이 안정적 수준에 도달했으며, 오염과 범죄의 경우 그 방향이 역전됨에 따라 최근에는 삶의 질을 오히려 향상시킨 것으로 본다. 나머지 부정적인 요인들은 중요하지만, 상품·서비스의 품질 및 다양성과 상점의 선택범위가 우리의 바이어스 척도에 부분적으로 반영된, 그러나 전적으로 포착되지는 못한 차원으로 증가하였으며, 그리고 가장 중요하게는 내구성의 엄청난 증가와 함께 아마 다른 것들을 모두 압도하면서 상쇄되고도 남았을 것으로 보인다. 따라서, CPI의 현재 바이어스에 대한 우리의 추정치가 만약 현실과 맞지 않는다면, 아마도 과소 평가되었을 가능성이 높다.

## Ⅷ. 위원회의 권고안

### 도입

우리의 권고안은 다음 집단을 대상으로 한 것이다 : 1. 선출직 공직자(대통령과 의회), 이들은 우리의 후원인이며 감독관인 동시에 소비자물가지수 정보의 주된 이용자이다; 2. 노동통계국, 소비자물가지수의 작성주체이다; 3. 전문 경제 및 통계학자, 이들은 궁극적으로 우리 경제의 주요 측정 및 관측 도구를 성공적으로 운용하기 위한 인력과 지식기반을 제공한다. 노동통계국은 주요 국가(경제)관측기관 중 하나를 운영하고, 경제상황에 관한 시의성 있는 정보를 생산하며, 아울러 경제가 기능하고 있는 방식 및 경제의 다양한 부문간 상호관계에 대하여 우리의 지식증진 및 이해심화를 도모할 수 있는 정보를 제공하고 있다.

### 권고안

1. 노동통계국은 소비자물가를 통해 달성하고자 하는 목적이 생계비 지수임을 명확히 해야 한다.

우리의 모든 개별 권고는 이를 위한 것이다.

2. 우리는 두 가지 지수를 개발하여 공표할 것을 권고한다.

첫째 지수는 매월 공표하고, 생계비지수의 정신은 유지하되 이에 필요한 정보가 때로는 일관성을 결여할 수 있음을 고려한다. 두 번째 지수는 연간으로 공표 및 업데이트하고, 새로운 정보와 연구결과를 도입하

기 위해 소급 개정한다. 두 개의 지수를 작성하는 목적은 복잡한 이슈들을 처리할 필요성 및 필요한 데이터 확보가 지연되는 문제에 대응하는 데에 있다.

노동통계국에 대한 권고는 3부분으로 구성된다 : 1. 단기: 약간의 추가적인 자원 또는 자료수집 활동으로 즉시 해결이 가능한 것이다. 첫째, 보다 최근 현황을 반영하도록 현행 소비자물가지수의 작성방법을 변경하고, 둘째, 매년 업데이트하며 이에 따라 개편되는 생계비지수를 산출하는 데에 중점을 두고 있다; 2. 중기: 현재의 기술수준에서 실행 가능한 개혁이나, 신규데이터 수집, 조직개편, 그리고 소비자물가지수를 통해 산출되는 다양한 하위지수의 세부내용 변경 등이 요구된다; 3. 장기적 권고: 추가 연구와 주의집중이 요구되는 주제 및 영역들에 초점을 두고 있다.

## 단기 권고안

3. 월간지수는 소비자물가지수의 명칭은 계속 유지하며, 장바구니 변화를 반영하는 “Superlative” 지수 산식을 채택하고 라스파이레스 산식을 고수하겠다는 구실을 포기함으로써, 생계비지수의 개념으로 이행해야 한다.

최하위 그룹에서 품목가격 변화를 집계하는 데에 필요한 수량정보의 확보가 지연되므로, 노동통계국은 소비자들이 밀접한 대체재의 가격변동에도 전혀 반응하지 않는다는 가정을 포기해야 한다. 따라서 우리는 노동통계국이 층(stratum) 및 기장수준품목(Entry level items, ELI)에 “trailing Tornqvist” 지수(가격비율의 가중기하평균)를 도입하고, 동시에 기초집계

수준에도 상대가격의 기하평균을 도입하도록 권장한다. 두 가지 방법 모두, 10년 전 소비패턴에 근거한 "바구니"의 관련성 저하문제를 경감시키고, 대체바이어스와 산식바이어스를 현저히 감소시키며, 아울러 신규상품과 서비스가 지수에 신속히 도입되도록 촉진할 것이다.

A. 소비자 지출패턴에 관한 최신정보 수집이 지연되므로, 과거 2년 또는 3년간 지출 연속평균값을 기초로 가중치를 산출해야 할 것이다. 즉 1996년의 가격변동에 대해서는 1993-94년의 가중치를 사용한다. 가중치는 매년 변경해야 한다.

B. 이는 소비자물가지수를 10년 주기가 아니라 상시 개편할 수 있도록 노동통계국 조직을 정비해야 함을 의미한다. 가중치와 조사상품 및 서비스 모두 보다 빈번하게 업데이트되어야 한다.

C. 또한 가능하다면 스캐너 데이터와 기타 "외부"데이터를 사용하여 데이터 수집비용을 절감하고, (우선) 가격이 조사되는 상품과 서비스를 확대하며, 현시점의 품목 가중치를 산출하고, 그리고 신규품목이 시장에 등장하는 즉시 반영되도록 한다. 이것이 데이터 수집비용의 순감소를 가져올 것인지는 단언하기 어렵다.

D. 공표된 CPI에 포함된 정보는 신규정보가 허용하는 한 과거로 소급하여, "superlative-index" 산식을 이용한 새로운 연간생계비지수 형태로 개편해야 한다. 이는 차기데이터 이용이 가능하고, 가중치의 업데이트 및 신규상품의 도입과 그 변천이력에 대한 확장 유추작업이 진행되면서 가능한 일이다. "개편된" 생계비지수는 1년 또는 2년이 경과한 후 연간으로 공표하며, 새로운 정보의 출현 및 신규기법의 도입과 함께 추가로 개편될 것이다. 이 지수의 공표치가 반드시 '최종적'일 필요는 없다.

#### 4. 노동통계국은 기초집계 수준에서 기하평균으로 이행해야 한다.

우리는 기하평균이 생계비 지수개념을 완벽하게 적용하는 데에 가장 근접해 있는 것으로 판단하고 있다. 현재 표본 추출된 상품중 대부분에 있어서 스캐너 데이터가 이용 가능하게 되면서, 궁극적으로는 기초 집계 수준에서도 기하평균을 실시할 수 있을 것으로 생각한다. 기하평균으로 변경할 경우, “사전적 평활화(pre-seasoning)”로의 전환을 통해 최근에 상당부분 제거된 “산식 바이어스”문제를 해결할 뿐만 아니라, 층(stratum) 이하수준의 대체 바이어스 또한 경감시킬 것이다. 그러나 이 경우에도 ‘상점 바이어스’는 해결하지 못할 것이다. 현재 센서스로 조사하고 있으나 노동통계국에서는 접근이 불가능한 상품수준의 세부 출하자료를 노동통계국이 확보함으로써 필요한 가중치를 업데이트하는 데에 도움이 되도록 해야 한다.

#### 중기 권고안

5. 노동통계국은 지수 구성요소의 행태를 연구하여, 어떤 요소들이 개별지수의 향후 장기흐름에 대한 정보를 가장 많이 제공하는지, 그리고 어떤 품목이 총합계와는 거의 무관한 변동을 보이는지 확인한 후, 자료수집활동에서 전자에 중점을 두어야 한다.

이는 특정 도시의 자료수집 비중이 감소하거나 심지어는 자료수집 자체가 배제되는 결과를 가져올 수 있다. 그리고 일부 상품은 전국적인 시장을 갖는 것으로 보고, 표본 추출시 품목 수는 크게 하는 반면 지역은 거의 고려하지 않는 반면, 연료비, 임대료, 개인서비스 및 신선식품 등 보다 “지역적인” 상품에 한하여 지역간 격차에 초점을 맞추는 방향으로 지수의 상품구조를 개편할 수 있다.

현재 노동통계국은 소비자물가지수에 포함된 추세의 흐름과 체계적인 연관이 없음에도 불구하고, 자료수집이 저렴하고 가격변동이 빈번한 것을 이유로, 바나나가격시세를 상당수 수집하고 있다. 동시에 외과진료, 가전제품 및 정보통신서비스 등 가격변동은 덜 빈번하나 변화(소멸 또는 재설계)의 가능성이 보다 높고, 가격측정이 곤란한 상품에 대해서는 주의를 덜 기울이고 있다.

6. 노동통계국은 지역에 대한 강조를 줄이는 방향으로 소비자물가지수의 표본추출 절차를 변경해야 한다. 우선, 가격이 조사될 상품 모집단을 추출한 후, 어떤 방식으로 어떤 상점에서 수집하는 것이 대표성 있는 가격표본 수집을 위해 가장 효율적인지 각 상품별로 결정해야 한다. 그 후에 비로소 자료수집의 경제성을 위해 지리적으로 군집된 표본에 의존해야 한다.

현재의 도시수준 물가지수는 지역별 수준비교에 전혀 쓸모가 없다. 또한 이는 명확히 정의된 애초의 수준을 바탕으로 하고 있지 않으므로 변동률 측정치로도 오해를 불러일으키기 쉽다. 지역별 물가수준을 제대로 설명하기 위해서는, 동일한 상품과 서비스의 가격을 서로 다른 도시에서 조사할 필요가 있다. 시장경쟁의 정도 및 인구추세간 차이에 기인하는 지역간 물가변동 격차를 연구하기 위해서는, 1년에 한번 정도 세부도시에서 “전국적” 상품들을 교대로 조사하는 것으로 충분할 가능성이 있다. 보다 일반적으로는, 월별추정의 주된 관심사인 소비자물가지수에 포함된 “전국적” 추세수준과, 보다 낮은 빈도로 자료를 수집하며 보다 완만한 변화속도를 보이는 도시간 격차로 구성된 모형을 구축할 수 있다.

이를 통해 소비자물가지수는 표본을 확대하고, 상품 및 서비스의 스펙트럼 중 의료서비스, 정보통신서비스 및 외식과 같이 변화가 빠르고, 품질변화와 상품매상고가 지역별로 다른(endemic) 분야를 분석하는 데에 자원을 집중할 수 있다.

매년 새로운 ‘장바구니’로 이행하는 개념을 도입하면 신규품목 및 상점을 보다 신속하게 반영할 수 있을 것이다. 신규품목중 대부분에서 전국적 표본추출을 지향할 경우, 특정 기장수준 품목(ELI)중 표본으로 추출되는 세부품목(모델, 종류, 유형) 수를 확대하는 것이 가능하다. 그리고 이는 무리하게 대체되는 품목 감소를 가져올 것이다. 가격에 관한 스캐너 데이터 등 새로운 자료원천과, 세부 품목에 관한 (보다 세부적인 가중치를 위한) 산업전반의 매출정보 등을 이용하는 것도 가능하게 될 것이다. 그 외에도 신규상점 및 서비스형태, 즉 이전에 소비한 것과 동일하거나 유사한 품목들을 보다 저렴하게 구입할 수 있는 효과적인 수단이 등장했음을 인지하여, ELI수준에서 보다 광범위한 품질조정과 비교성 있는 품목대체를 수행할 수 있다.

**7. 노동통계국은 분류, 즉 품목군을 정의하는 기준이 물가지수에 미치는 파장을 조사함으로써, 지수가 품목대체를 전적으로 포착할 수 있도록 해야 한다.**

그뿐 아니라, 신규상품을 소비자들이 이로써 대체할 가능성이 가장 높은 상품들과 동일한 하위군(층, stratum)으로 분류하는 규칙을 적용해야 한다. 신문과 경쟁관계에 있는 온라인 뉴스서비스, 자동차 리스와 구입, 약품 및 이로 인해 대체되는 진료행위 등은 대체효과의 전모를 밝히기 위해 직접 비교할 필요가 있는 상품의 예이다.

8. 주의를 요하는 개념적 이슈가 몇 가지 더 있다. 자동차등 내구재의 가격을 연간 서비스가격으로 변경해야 한다. 이는 자가소유주택 가격에 관한 현행 처리방법과 동일선상에 있다. 또한 “보험”의 처리도 현재 사용되고 있는 보험수익에 기초한 사후적 측정수단(ex-post insurance profits based measure)보다는 사전적 소비자물가 측정수단(ex-ante consumer price measure)으로 이행할 필요가 있다.

9. 노동통계국에 외부 정보, 전문지식 및 연구결과의 도입을 위한 보다 영구적인 메커니즘이 필요하다.

본 위원회는 소비자물가지수의 다양한 측면들을 모두 충분히 심도 있게 조사할 만큼 지원 또는 시간을 갖지 못했다. 향후 유사한 단체가 또다시 임시로 구축되는 경우에도 이와 마찬가지로 일 것이다. 노동통계국 요청하에, 미국경제학회, NRC-NAS 또는 NBER 등과 같이 독립된 공공 전문가단체로 조직된, 보다 항구적인 기구를 구축해야 한다. 이 단체는 노동통계국과 협력하여 보다 근본적인 연구조사를 추진하고, 자료수집 및 추정의 다양한 대안적 접근방법들을 실험하기 위한 체제를 제공할 수 있다. 또한 통상적인 자료수집과정에서 발생하는 통계, 경제 및 최신 시장이슈와, 지수구성, 그리고 여기서 제안한 개선의 일부 시행에 관하여 일정한 범위의 전문가와 업계의 소견에 접근할 수 있는 보다 영구적인 통로가 될 것이다.

## 장기 권고안

10. 노동통계국은 소비자물가지수의 현행 “장바구니” 구조를 극복할 수 있도록 연구프로그램을 개발해야 한다.

장기적으로는 신규상품 및 서비스, 그리고 소비자의 활동무대이며 변화하고 있는 경제, 사회 및 환경적 풍토 등이 대형 이슈에 속한다. 이 프로그램은 새로운 의료기법과 정보통신기기로 인해 절약되는 시간의 가치와 아울러 수명연장 및 이에 연관된 품질, 근린상가의 폐쇄로 인해 보다 멀리 떨어진 신설 쇼핑센터를 이용할 경우 거리 측면에서 입는 손실, 그리고 범죄 증가, 새로운 질병 또는 조세제도 변화로 인해 소비자가 강요당하는 “소비”증가의 가치를 측정하는 방안을 연구해야 한다.

#### **11. 노동통계국은 균형가격의 가정으로부터 파생되는 효과를 조사해야 한다.**

이 가정은 가격 또는 수량이 품질변화 또는 대체재의 등장에 즉시 반응하는 것을 의미하고, 이 방법론의 요소들중 여러 개의 본질적인 전제이다. 이 가정이 때때로 성립하지 않는 것은 본 보고서에서 논의한 많은 이슈들의 핵심에 있다. 노동통계국은 연구가 허용되는 한, 이 주된 가정을 완화하기 위해 필요한 방법론적 변경이 무엇인지 확인할 것을 권고한다.

#### **12. 노동통계국은 이 방향으로 발전을 이룰 수 있도록 몇 가지 신규 데이터수집 이니셔티브를 개발할 필요가 있다.**

가장 우선으로, 노동통계국 또는 관련기관은 대규모 소비자표본을 통해 시간사용에 관한 세부자료를 수집할 필요가 있다. 또한 현재의 ‘보건상태조사’를 확장하여 다양한 “삶의 질” 이슈에 관련된 정보를 보다 많이 포함시켜야 한다. ‘범죄피해조사’ 및 물리적 환경상태의 다양한 척도로부터 자료를 통합하는 실험적 사회-경제 위성국민계정(satellite national account)의 측면에서도 진전이 이루어져야 한다(이 분야에서

환경위성계정 등을 이미 수행하고 있는 경제분석국(BEA)과 공동작업도 가능하다). 이 계정은 소비 장바구니뿐만 아니라 소비로 인해 평균적인 개인이 경험하는 여가와 삶의 질에 대해서도 가치를 평가한다. 다양한 연령 및 사회계층별 측정치분포에 관한 정보를 제공할 수도 있을 것이다. 이들을 생계비의 주된 구조 속으로 통합하는 일이 용이하지는 않겠지만, 향후 이 방면에서 진전을 통해 정책결정 담당자 및 일반대중에게 유용한 보충정보가 제공될 것임에 틀림없다.

## 의회에 대한 제언

13. 의회는 경제통계의 정확성과 시의성 증진을 염두에 둔 정보공유와 경제통계의 개발 및 생산에 소요되는 자원절감을 위해 상무부와 노동부에게 필요한 법령을 제정해야 한다.

14. 의회는 소비지출조사(CES)의 표본과 수집항목을 확대하고, 구매시점조사(POPS)의 빈도를 높이는 동시에 산업조사 및 스캐너 데이터 등 여타 국내 원천으로부터 상품 상세정보를 보다 많이 확보하기 위해 필요한 자원을 추가로 제공해야 한다.

위원회는 향후 비용절감의 가능성이 있는 분야가 무엇인지 확인했으며, 민간사업체 부문의 기법들을 이용하여 연방정부의 효율성을 제고하는 최근 추세에 공감하고 있다. 반면 사업체들이 하드웨어 등의 유형자산에서 점차 중요성을 더해 가는 소프트웨어 등의 무형자산, 그리고 인적자산에 이르기까지 정보기술에 막대한 투자를 하는 압도적인 추세 또한 주목하고 있다.

15. 의회는 이 분야의 발전을 3년여 주기로 점검하고 통계현황에 대해 의회에게 적절한 해설을 자문하기 위해 상설(순환) 독립위원회 또는 전문위원단을 설립해야 한다.

이는 그 자체로서 유용하나, 특히 새로운 지수로의 이행을 수월하게 하는 측면에서 유용하다.

16. 의회와 대통령은 다양한 연방지출 프로그램 및 조세법령 규정에 널리 퍼져있는, 현저한 과대 물가연동이 지속되기를 바라는지 결정해야 한다. 만약 물가연동의 목적이 이전지출 수령집단과 납세 집단을 완전하고 정확하게 분리하는 데에 있으며, 그 이상도 그 이하도 아니라면, 의회는 물가연동 규정을 이에 맞도록 조정한 법령을 통과시켜야 한다.

이는 현행 소비자물가지수에 연동된 값으로부터 과다연동을 부분적으로 또는 전적으로 반영한 값을 공제한다는 맥락에서 행할 수 있다. 다른 방안으로, 새로 개편된 연간지수를 개발하여 정기적으로 공표할 경우, 동 지수에서는 이보다 작은 값을 공제함으로써, 생계비변동에 보다 근접할 필요가 있을 것이다.

물가연동 프로그램에 보다 정확한 생계비지표라는 협소한 범위를 초월한 다른 측면들이 여럿 있으며, 그 외에 이슈들도 다수 야기됨을 서둘러 덧붙인다. 또한 노동통계국, 의회, 행정부(executive branch), 그리고 국민들이 이 견해 및 그 함의를 완전히 숙고하고 또 이해할 필요가 있음을 표명하고자 한다.

## 경제학자와 통계학자들에게

이들은 자료수집, 자료분석 및 해석 훈련을 보다 중요시하고, 정규 커리큘럼중 보다 많은 시간과 주의집중을 기울여야 한다. 경제학자와 통계학자를 양성하는 데에 있어서 측정 및 표본추출 이슈를 더욱 강조할 필요가 있다. 정부내의 전문가들과 학계 및 업계의 동료들 사이 유대를 개선하기 위한 노력 또한 기울여야 한다. 학계에서는 자신의 뒤이은 분석에 필요한 “원시자료”를 대부분 제공하는 정부내의 동료들이 중요한 업무를 수행하고 있음을 인식하고, 그 노고에 대한 감사 및 그들의 제한된 근무여건에 대한 이해심을 표해야 한다.

## IX. 결론

CPI는 현재 이용 가능한 것 중에서는 최고의 척도임에도 불구하고, 진정한 생계비지표는 아니다. 생계비변동의 측정수단으로서 CPI는 다양한 개념적인, 그리고 현실적인 문제점을 안고 있다. 노동통계국이 CPI에 중대한 업데이트 및 개선을 가했음에도 불구하고, CPI 변동폭이 실제 인플레이션율을 상당히 과대평가 해왔을 가능성이 높다. 그뿐 아니라, 개정은 시간의 연장선상에서 내부적으로 일치되는 일련의 생계비 값이 제공되도록 이루어지지 않았다. 보다 중요하게, 소비자물가지수 변동은 진정한 생계비변동을 가까운 장래에 계속 과대평가할 가능성이 높다. 이 과대평가는 정부지출 및 조세구간의 과대연동과 연방 재정적자 및 부채의 증가를 포함하여, 의도하지 않은 중대한 결과를 초래할 것이다. 만약 물가연동의 취지가 이전소득 수혜자 및 납세자를 생계비변동으로부터 보호하는 데에 있다면, 소비자물가지수를 사용함으로써, 과거와 마찬가지로 미래에도 이들은 진정한 생계비변동에 비해 (평균적으로) 과다한 보상을 받게 될 것이다.

본 보고서는 보다 정확한 생계비척도를 개발하는 데에 있어서 다루어야 하는 다양한 이슈들을 나열하였다. 본 보고서는 또한 미국 물가통계 담당기관과 이 통계들을 후원, 감독, 그리고 이용하는 선출직 공직자들에 대한 일련의 권고를 제시하였다. 우리의 권고를 이행할 경우 미국 물가통계의 정확성이 현저히 개선될 것임을 믿어 의심치 않는다. 이는 민간계약 및 공공 프로그램에 대한 생계비조정에서, 미국연방 준비은행에 대한 국민소득 및 국민생산계정의 구성요소 개선 정보제공에 이르기까지 모든 것을 보다 정확히 측정하도록 할 것이다. 이 진전은 다시 시민과 정책결정 담당자들이 모두, 시간에 걸쳐, 여러 그룹 간, 그리고 각 국가간 경제발전을 측정하는 데에 도움을 줄 것이다.

소비자물가지수 변동을 기초로 한 생계비변동의 과대평가에 관한 위원회의 최선의 추정치는 1%를 약간 초과하는 데에 비해, 보다 넓은 관점에서 볼 경우 미세한 격차도 시간이 지남에 따라 축적되고 중대하다는 견해이다. 이는 1980년 초반 자가소유 주택의 처리방법을 개선한 사례를 통해 제Ⅱ부에서 입증되었다. 최근 산식 바비어스를 정정함으로써 1979-1995기간에 매년 0.24%의 바이어스가 추가된 경우도 이에 해당한다. 누적적인 파생결과는 현저하다.

뒤이은 분석, 연구, 그리고 경제 상황에 의해 이 추정치가 약간 변경될 가능성이 있는 한, - 상향 또는 하향으로 변화가 일어날 가능성은 마찬가지로인 것으로 생각한다 - 그 사용에 있어서 어느 정도 신중함이 마땅하다. 본 보고서에 제시된 분석이 우리의 최선의 판단을 나타냄에도 불구하고, 미국 물가통계의 점검을 위한 이전의 주된 시도인 소위 Stigler 위원회가 1961년 분석당시 보유했던 상당한 자료원천을 본 위원회는 갖지 못했다. Stigler 위원회는 상당한 시초연구를 의뢰 및 수행할 수 있었던 반면, 본 위원회에게는 이를 위한 시간과 자원이 주어

지지 않았다. 그럼에도 불구하고, 본 보고서는 정부내외 모두에 걸친 광범위한 원천으로부터 받아들인 새로운 정보를 구체화하고 있다. 우리는 위원회의 조사과정에서 야기된 개별 이슈들에 대하여 엄청난 제의, 조언, 및 평가들이 밀려들어온 데에 감사한다.

독자들은 보고서의 결과, 분석 및 권고를 숙고하고 이해할 시간이 필요하다. 여기에는 노동통계국, 의회, 행정부, 민간부문 그리고 학계가 모두 포함된다. 본 보고서에 제시된 견해들을 본래의 정신, 즉 『미국 물가통계의 작성 및 사용을 개선하는 방법에 관한 몇 가지 지침과, 복잡하고 역동적인 경제 하에서 물가지수가 끊임없이 개선되는 과정을 제시하기 위한 하나의 시도』에 의거하여 주의 깊게, 그리고 세심하게 검토하길 진심으로 바란다.



## 원문 : The Boskin Commission Report

Towards a More Accurate Measure Of The Cost Of Living



## **Executive Summary**

1. The American economy is flexible and dynamic. New products are being introduced all the time and existing ones improved, while others leave the market. The relative prices of different goods and services change frequently, in response to changes in income and technological and other factors affecting costs and quality. This makes constructing an accurate cost of living index more difficult than in a static economy.

2. Estimating a cost of living index requires assumptions, methodology, data gathering and index number construction. Biases can come from any of these areas. The strength of the CPI is in the underlying simplicity of its concept: pricing a fixed (but representative) market basket of goods and services over time. Its weakness follows from the same conception: the "fixed basket" becomes less and less representative over time as consumers respond to price changes and new choices.

3. There are several categories or types of potential bias in using changes in the CPI as a measure of the change in the cost of living. 1) Substitution bias occurs because a fixed market basket fails to reflect the fact that consumers substitute relatively less for more expensive goods when relative prices change. 2) Outlet substitution bias occurs when shifts to lower price outlets are not properly handled. 3) Quality change bias occurs when improvements in the quality of products, such as greater energy efficiency or less need for repair, are measured inaccurately or not at all. 4) New product bias occurs when new products are not introduced in the market basket, or included only with a long lag.

4. While the CPI is the best measure currently available, it is not a true cost of living index (this has been recognized by the Bureau of Labor Statistics for many years). Despite many important BLS updates and improvements in the CPI, changes in the CPI will overstate changes in the true cost of living for the next few years. The Commission's best estimate of the size of the upward bias looking forward is 1.1 percentage points per year. The range of plausible values is 0.8 to 1.6 percentage points per year.

5. Changes in the CPI have substantially overstated the actual rate of price inflation, by about 1.3 percentage points per annum prior to 1996 (the extra 0.2 percentage point is due to a problem called formula bias inadvertently introduced in 1978 and fixed this year). It is likely that a large bias also occurred looking back over at least the last couple of decades.

6. The upward bias creates in the federal budget an annual automatic real increase in indexed benefits and a real tax cut. CBO estimates that if the change in the CPI overstated the change in the cost of living by an average of 1.1 percentage points per year over the next decade, this bias would contribute about \$148 billion to the deficit in 2006 and \$691 billion to the national debt by then. The bias alone would be the fourth largest federal program, after social security, health care and defense. By 2008, these totals reach \$202 billion and \$1.07 trillion, respectively.

7. Some have suggested that different groups in the population are likely to experience faster or slower growth in their cost of living than recorded by changes in the CPI. We find no compelling evidence of this to date (in fact just the opposite) but further exploration of this issue is desirable.

8. The commission is making over a dozen specific recommendations to the BLS. These include the following:

i. The BLS should establish a cost of living index (COLI) as its objective in measuring consumer prices.

ii. The BLS should develop and publish two indexes: one published monthly and one published and updated annually and revised historically.

iii. The timely, monthly index should continue to be called the CPI and should move toward a COLI concept by adopting a "superlative" index formula to account for changing market baskets, abandoning the pretense of sustaining the fixed-weight Laspeyres formula.

iv. The new annual COL index would use a compatible "superlative-index" formula and reflect subsequent data, updated weights, and the introduction of new goods (with their history extended backward).

v. The BLS should change its procedure for combining price quotations by moving to geometric means at the elementary aggregates level.

vi. The BLS should study the behavior of the individual components of the index to ascertain which components provide most information on the future longer-term movements in the index and which items have fluctuations which are largely unrelated to the total and emphasize the former in its data collection activities.

vii. The BLS should change the CPI sampling procedures to de-emphasize geography, starting first with sampling the universe of commodities to be priced and then deciding, commodity by commodity, what is the most efficient way to collect a representative sample of prices from which outlets, and only later turn to geographically clustered samples for the economy of data collection.

viii. The BLS should investigate the impact of classification, that is item group definition and structure, on the price indexes to improve the ability of the index to fully capture item substitution.

ix. There are a number of additional conceptual issues that require attention. The price of durables, such as cars, should be converted to a price of annual services, along the

same lines as the current treatment of the price of owner-occupied housing. Also, the treatment of "insurance" should move to an ex-ante consumer price measure rather than the currently used ex-post insurance profits based measure.

x. The BLS needs a permanent mechanism for bringing outside information, expertise, and research results to it. At the request of the BLS, this group should be organized by an independent public professional entity and would provide BLS an improved channel to access professional and business opinion on statistical, economic and current market issues.

xi. The BLS should develop a research program to look beyond its current "market basket" framework for the CPI.

xii. The BLS should investigate the ramifications of the embedded assumption of price equilibrium and the implications of it sometimes not holding.

xiii. The BLS will require a number of new data collection initiatives to make some progress along these lines. Most important, data on detailed time use from a large sample of consumers must be developed.

9. The Commission is making several recommendations to the President and Congress. These include the following:

xiv. Congress should enact the legislation necessary for the Departments of Commerce and Labor to share information in the interest of improving accuracy and timeliness of economic statistics and to reduce the resources consumed in their development and production.

xv. Congress should provide the additional resources necessary to expand the CES sample and the detail collected, to make the POPS survey more frequent, and to acquire additional commodity detail from alternative national sources, such as industry surveys and scanner data.

xvi. Congress should establish a permanent (rotating) independent committee or commission of experts to review progress in this area every three years or so and advise it on the appropriate interpretation of then current statistics.

xvii. Congress and the President must decide whether they wish to continue the widespread substantial overindexing of various federal spending programs and features of the tax code. If the purpose of indexing is accurately and fully to insulate the groups receiving transfer payments and paying taxes, no more and no less, they should pass legislation adjusting indexing provisions accordingly.

This could be done in the context of subtracting an amount partly or wholly reflecting the overindexing from the current CPI-based indexing. Alternatively, a smaller amount would need to be subtracted from indexing based on the new revised annual index if and when it is developed and published regularly, to more closely approximate the change in the cost of living.

We hasten to add that the indexed programs have many other features and raise many other issues beyond the narrow scope of a more accurate cost of living index. We also wish to express our view that these findings and their implications need to be fully digested and understood by the BLS, the Congress, the Executive Branch and the public.

## **I. Introduction<sup>1</sup>**

Accurate measures of changes in the cost of living are among the most useful and important data necessary to evaluate economic performance. The change in the cost of living between two periods, for example 1975 and 1995, tells us how much income people would have needed in 1975, given the prices of goods and services available in that year, to be at least as well off as they are in 1995 given their income and the prices of goods and services available then. For example, if a family with a \$45,000 income in 1996 would have needed \$15,000 in 1976, the cost of living has tripled in the interim.

If the American economy was quite static, with very few new products introduced, very little quality improvement in existing products, little change in consumers' income, and very small and infrequent changes in the relative prices of goods and services, measuring changes in the cost of living would be conceptually quite easy and its implementation a matter of technical detail and appropriate execution. Fortunately for the overwhelming majority of Americans, our economy is far more dynamic and flexible than that. New products are being introduced all the time and existing ones improved, while others leave the market. The relative prices of different goods and services change frequently, in response to changes in consumer demand, and technological and other factors affecting costs and quality. Consumers in America have the benefit of a vast and growing array of goods and services from which to choose, unlike consumers in some other countries or our ancestors many decades ago. But because the economy is complex and dynamic is no reason to bemoan the greater difficulty in constructing an accurate cost of living index. Major improvements can and should be made to the various official statistics that are currently used as proxies for changes in the cost of living, such as the well-known Consumer Price Index (CPI). The Consumer Price Index measures the cost of purchasing a fixed market basket of goods and services. Based on surveys of households from some base period, the index sets weights (expenditure shares) for different goods and services. The weights reflect average or representative shares for the groups surveyed. <sup>2</sup> Keeping these weights fixed through time, the CPI is then calculated by attempting to measure changes from one month to the next in prices of the same, or quite closely related, goods and services.

But through time consumption baskets change, in part because of changes in the relative prices of goods and services, and therefore the weights from the base period no longer reflect what consumers are actually purchasing. Representative purchases also change as discount coupons, buyers' clubs and other marketing devices determine the best value and alter buying patterns. This failure to adjust for the changes in consumer behavior in response to relative price changes is called substitution bias. It is a necessary result of keeping the market basket fixed. Because the market basket is updated only every decade or so, as we get further away from the base period, there is

more opportunity for relative prices to diverge from what they were in the base period, and for consumption baskets to change substantially.

Just as there are changes in what consumers purchase, there are also trends and changes in where purchases are made. In recent years, there has been a transformation of retailing. Superstores, discount stores and the like now comprise a large and growing fraction of sales relative to a decade or two ago. As important as keeping up with the basket of goods that consumers actually purchase is keeping up with the outlets where they actually purchase them, so that the prices paid are accurately recorded. The current methodology suffers from an outlet substitution bias, which insufficiently takes into account the shift to discount outlets.

Many of the products sold today are dramatic improvements over their counterparts from years ago. They may be more durable and subject to less need for repair; more energy efficient; lighter; safer; etc. Sometimes, at least initially, a better quality product replacing its counterpart may cost more. Separating out how much of the price increase is due to quality change rather than actual inflation in the price of a standardized product is far from simple, but is necessary to obtain an accurate measure of the true increase in the cost of living. To the extent quality change is measured inaccurately or not at all, there is a quality change bias in the CPI.

The same is true with the introduction of new products, which have substantial value in and of themselves -- not many of us would like to surrender our microwave ovens, radial tires, and VCR's -- as well as the value of greater choice and opportunities opened up by the new products. To the extent new products are not included in the market basket, or included only with a long lag, there is a new product bias in the CPI. Finally, in a dynamic, complex economy like the contemporary United States, there are literally many thousands of goods and services consumed. Price data are collected at a considerable level of disaggregation and how the price changes are aggregated into an overall index involves quite technical issues that can lead to a formula bias in the CPI.

Even if no federal program on either the outlay or revenue side of the budget were indexed, it would still be desirable to improve the quality of measures of the cost of living from the standpoint of providing citizens a better and more accurate estimate of what was actually going on in the economy, a way to compare current performance to our historical performance or to that of other countries. For example, the most commonly used measure of the standard of living is real income or output per person. To measure changes in real income requires the separation of nominal income changes from price changes. Obviously, that requires an accurate measure of price changes. The Commerce Department uses the component indexes of the CPI as inputs in estimating inflation and real GDP, and thus some of the bias from the CPI is transmitted to the national income accounts.

But numerous federal, state and local government programs and tax features are "indexed" for changes in the cost of living by the changes in the Consumer Price Index. The CPI is also used to index, formally or informally, a large number of private sector contracts, including wages in collective bargaining agreements and rents, to name obvious examples that affect millions of Americans. Currently, slightly under one-third of total federal outlays, mostly in retirement programs, are directly indexed to changes in consumer prices. Several features of the individual income tax, including the tax brackets, are indexed; the individual income tax accounts for a little under half of federal revenues.

Congress indexed these outlay programs and tax rules in order to help insulate or protect the affected individuals from bearing the brunt of increases in the cost of living. Yet the Bureau of Labor Statistics, the agency responsible for compiling and presenting the Consumer Price Index, has explicitly stated for years that the CPI is not a cost of living index, presumably for some of the reasons mentioned above. If the Consumer Price Index as currently produced, and as likely to be produced over the next few years, is not an appropriate cost of living index for the task Congress had in mind, then it is desirable to consider alternative measures.

The consequences of changes in the Consumer Price Index overstating changes in the cost of living can be dramatic. For example, if use of the CPI is expected to overstate the increase in the cost of living by one percentage point per year over the next dozen years, the national debt would be about \$1 trillion greater in 2008 than if a corresponding correction were made in the indexing of outlays and revenues.

This report proceeds as follows: Section II discusses the historical and prospective budgetary implications of changes in the CPI overstating changes in the cost of living. Section III presents an overview of how the CPI is actually constructed. Section IV details why the CPI is not a true cost of living index and discusses substitution bias. Section V describes in greater detail the current procedures employed by the BLS to adjust for quality change and presents a survey of the studies and the Commission's judgment on the bias from quality change and new products. Section VI summarizes the Commission's findings on the size of the bias by type, plus the range of plausible overall bias. Section VII discusses the issue of separate price indexes for different groups and of aspects of the quality of life that fall primarily outside the market based consumption focus of cost-of-living measures. Section VIII presents the Commission's detailed recommendations of ways to produce and to use more accurate cost-of-living measures. The Conclusion offers a brief perspective and some cautionary notes on the use of the findings of the Commission.

## **II. Indexing The Federal Budget**

The issue posed for fiscal policy makers by an upward bias in the CPI has been stated with admirable clarity by the Congressional Budget Office (1994):

The budgetary effect of any overestimate of changes in the cost of living highlights the possibility of a shift in the distribution of wealth. If the CPI has an upward bias, some federal programs would overcompensate for the effect of price changes on living standards, and wealth would be transferred from younger and future generations to current recipients of indexed federal programs -- an effect that legislators may not have intended. <sup>5</sup>

Social Security is by far the most important of the federal outlays that are indexed to the CPI. However, Supplemental Security Income, Military Retirement, and Civil Service Retirement are significant programs that are similarly indexed. Other federal retirement programs, Railroad Retirement, veterans' compensation and pensions, and the Federal Employees' Compensation Act also contain provisions for indexing. The Economic Recovery Tax Act of 1981 indexed individual income tax brackets and the personal exemption to the CPI.

How important have the budgetary consequences of upward bias in the CPI been historically? Obviously, a precise answer to this question would require extended study, taking into account the timing of the bias, the parallel development of indexing provisions in specific federal outlays and revenues, and interest on the accumulation of debt that has resulted. An indication of the potential size of these effects can be inferred from one important historical example of one clearly identified source of bias. A careful study of this type, which focuses on the most important federal program affected by indexing, namely, social security benefits, has been conducted by the Office of Economic Policy (OEP) of the Department of the Treasury.

On February 25, 1983, the Bureau of Labor Statistics (BLS) introduced an important technical modification in the Consumer Price Index for All Urban Consumers (CPI-U). This altered the treatment of housing costs by shifting the costs for homeowners to a rental equivalent basis. The new treatment of housing costs was incorporated into the Consumer Price Index for Urban Wage Earners and Clerical Workers (CPI-W), used to index social security benefits, in 1985.

The rental equivalent measure of housing costs was a conceptual improvement and has been retained in subsequent official publications. However, housing costs in preceding years employed a "homeownership" measure " . . . based on house prices, mortgage interest rates, property taxes and insurance, and maintenance costs." <sup>4</sup> The treatment of housing costs prior to 1983 was not modified in publishing the revised CPI-U, so that the new treatment of housing introduced a discrepancy in the conceptual basis for the CPI-U before and after 1983. Similarly, housing costs in the CPI-W prior to 1985

have not been modified.

BLS developed an "experimental" price index, CPI-U X1, based on a rental equivalent treatment of housing extending back to 1967. This provides the basis for the OEP assessment of bias in the CPI-W. The bias for 1975, the first year that social security was indexed to the CPI-W, was 1.1 percent. This bias mounted over subsequent years, reaching 6.5 percent by 1982 and then declining to 4.7 percent in 1984. <sup>3</sup>

Overpayments of social security benefits resulting from the bias in the CPI-W mounted through 1983, reaching a total of \$8.76 billion or 5.55 percent of benefits paid in that year. These overpayments have resulted in a lower balance in the OASI trust fund and a larger federal deficit and debt. OEP estimates interest costs associated with these deficits at the rate of interest paid or projected to be paid on the OASI trust fund. Beginning in 1984 interest costs predominate in the total. In the current fiscal year the total cost is \$21.79 billion, of which \$17.64 billion is interest. The cumulative effect of just this one source of bias in the CPI-W via this one program on the federal debt amounts to \$271.0 billion, as of 1996.

In summary, the BLS made two decisions in revising the treatment of housing costs in the CPI-W in 1985. The first decision was to change the treatment of housing costs to a rental equivalent basis beginning in January 1985. The second was not to revise the treatment of housing costs for 1984 and earlier years. As a consequence of these two decisions the level of the CPI-W is 4.7 percent above the CPI-U X1, a measure of the cost of living based on the same primary data sources and similar methodology, but with a consistent treatment of housing costs.

The increases in federal outlays resulting from the bias in the CPI-W cannot be justified as cost of living adjustments. These increases are the consequence of an inappropriate treatment of housing costs before 1985 and have resulted in large transfers to beneficiaries of the OASI program that are devoid of any economic rationale. The overpayments have continued up to the present, but are declining in importance. However, the resulting decline in the OASI trust fund continues to mount

due to rising interest costs and now contributes more than two hundred billion dollars to the federal debt.

Of course, nobody would suggest retroactively undoing the overindexing due to this or any other source of bias. The point of this discussion is to demonstrate how important it is to correct biases in the CPI as quickly and fully as possible before their consequences mount, indeed compound.

What would be the effect of an upward bias in the CPI on future budget deficits? More than half of federal spending of \$1.5 trillion is now attributable to entitlements and mandatory spending programs. In January 1995 the annual Congressional Budget Office (CBO) outlook for the economy and the federal budget showed that this proportion is projected to rise to almost two-thirds of federal spending during fiscal year 1998. Cost-of-living adjustments at a projected rate of 3.0 percent will contribute \$43 billion to total spending on mandatory programs in that year and \$80 billion in fiscal year 2000.<sup>6</sup> This is 6.8 percent of projected spending on mandatory programs in fiscal year 2000.

Testimony presented by the CBO to the Committee on Finance shows the impact of a hypothetical correction (reduction) of 0.5 percentage point in cost of living adjustments for fiscal years 1996-2000. Federal outlays would decline by \$13.3 billion in fiscal year 2000, while revenues would rise by \$9.6 billion. The decline in debt service resulting from reduced deficits in fiscal years 1996-2000 would be \$3.3 billion, yielding a total contribution to deficit reduction of \$26.2 billion in fiscal year 2000.<sup>7</sup> This is more than ten percent of the deficit projected by CBO in that year.

The CBO has provided the Commission with updated projections of the impact of hypothetical corrections (reductions) of 0.5 and 1.0 percentage point in cost of living adjustments for fiscal years 1997-2006.<sup>8</sup> With a reduction of 0.5 percentage point the total contribution to deficit reduction rises to \$67.5 billion in 2006. Of this amount, an increase in revenue accounts for \$22.3 billion and reductions in outlays, including debt service, amounts to \$45.3 billion (of which debt service is \$13.1 billion).

CBO projections for the impact of a hypothetical correction (reduction) in cost of living adjustments of 1.0 percentage point are, of course, even more dramatic. The total change in the deficit in the year 2006 is \$134.9 billion. Federal revenues would be increased by \$44.5 billion and federal outlays reduced by \$90.5 billion; of the reduction in outlays \$26.1 billion can be attributed to lower debt service and \$64.4 billion to lower outlays on indexed programs. (See Appendix Figure A-1 for detail).

Stated differently, if the change in the CPI overstated the change in the cost of living by an average of one percentage point per year over this period, this bias alone would contribute almost \$135 billion to the deficit in the year 2006. That is one-third the projected baseline deficit (which assumes no policy changes such as the current balanced budget proposals). More remarkably, the upward bias by itself would constitute the fourth largest federal outlay program, behind only social security, health care and defense. By 2008, the increased deficit would be \$180 billion and national debt \$1 trillion.

In summary, an upward bias in the CPI would result in substantial overpayments to the beneficiaries of federal entitlements and mandatory spending programs. In addition, such a bias would reduce federal revenues by overindexing the individual income tax. In short, the upward bias programs into the federal budget every year an automatic, real increase in indexed benefits and a real tax cut. Correction of biases in the CPI, while designed to adjust benefits and taxes for true changes in the cost of living more accurately, would also contribute importantly to reductions in future federal budget deficits and the national debt. These reductions can be attributed to higher revenues, lower outlays, and less debt service. Lower outlays -- cuts in indexed federal spending programs and reduced interest payments -- account for over two-thirds of the long-run deficit reduction, while higher revenues account for the rest.

### **III. How The CPI Is Constructed: A Brief Introduction**

Knowledge of how the CPI is constructed is needed to understand the reasons that biases occur and the rationale for our recommendations for improvements and

changes. This section provides a brief description of the BLS methodology highlighting the places where biases and key issues are likely to arise. We refer the reader to BLS documentation for more detail on data collection procedures and index construction methodology, as well as to recent articles by Armknecht (1996), and Shapiro and Wilcox (1996b).<sup>9</sup>

As could be inferred from the discussion above about the complexities of a modern dynamic capitalist economy, the CPI program is a complex and difficult undertaking. To make it manageable, the BLS applies a simplified view of the marketplace and consumer behavior. This simplified view is reflected throughout the CPI approach. It takes expenditures for a fixed market basket of goods and services at some point in the past, called the base or reference period, and estimates what it would cost today to purchase the same market basket. The formula used to construct the CPI, called Laspeyres, assumes that purchases are made in fixed quantities based on decisions from some previous period's experience. In other words, the CPI attempts to answer the question, "what is the cost, at this month's market prices, of purchasing the same market basket actually purchased in the base period?" Since the Laspeyres formula does not allow for the substitution of products or services in response to current prices and choices, it is an "upper bound" to a cost of living.

The market basket consists of total expenditures on items directly purchased by all urban consumers, that is, food, clothing, shelter and fuels, transportation, medical services and other goods and services that people buy for day-to-day living. The BLS uses scientific sampling techniques to select specific items. The BLS measures the price changes in these items over time. The sample design involves a multistage process for sampling by geographic area, retail outlet, item category, and individual goods and services within an outlet and category.

Several samples are used to try to make the CPI representative of the prices paid by consumers: urban areas selected from all U.S. urban areas, consumer units within each selected area, outlets from which these consumer units purchased goods and services, specific items -- goods and services -- purchased by these consumer units, and housing

units in each urban area (for the shelter component of the CPI). The key sources of information used to determine the items which comprise the market basket and the outlets at which prices are to be collected are the Consumer Expenditure Survey (CES) and the Point-of-Purchase Survey (POPS).

Each month, prices for approximately 71,000 goods and services are collected from 22,000 outlets, in 44 geographic areas. <sup>10</sup> Separately, information is collected each month from about 5,000 renters and 1,000 homeowners for the housing components of the CPI. <sup>11</sup> The price quotations are combined, that is, aggregated, into the overall CPI. The determination of representative items to be priced, the procedure for collecting prices at the outlets, and the levels at which the prices are combined into indexes and the indexes are combined into higher aggregates, are all based on a fixed structure or system in which a number of key assumptions are embedded.

The item structure has four levels of classification beginning with major groups such as food and beverages, transportation, and medical care. The seven major groups are made up of 69 expenditure classes (EC's), for example fresh fruits (EC 11) and hospital and other related services (EC 57). The expenditure classes are in turn divided into 207 groupings called item strata, the lowest level at which indexes are constructed. <sup>12</sup> Two examples of item strata are apples, and nursing and convalescent home care. It is important to note that while the item categories are mutually exclusive and exhaustive for all consumer expenditures, this does not mean that new goods and services are automatically brought into the sample if they were not available during the reference period. It just means that every good and service can be classified within an existing stratum and there is no need to create a new stratum for when a new good or service is introduced. (This is made possible in part by numerous item categories called "other.")

Within each item stratum, entry level items, called ELIs for short, are defined. Indexes are not constructed at this level. Many stratum have only one ELI for example, Apples. The ELIs are the lowest level sampling units for items. They are the level of item definition at which the data collectors begin item sampling within each sample outlet.

For example, prices for Brand "X" fever thermometers for babies, model 41303 41082, 4 3/10 inches long with plastic case, sold by "Y" Foods, Inc. in West Terre Haute Indiana, might be collected for Medical Equipment for General Use, ELI 55032, within Non-prescription Drugs and Medical Supplies, item stratum 5503, within Non-prescription Drugs, EC 55, within the Medical Care Commodities component of the Medical Care major expenditure category. <sup>13</sup> —

Outlets within geographic areas are sampled, too. The probability of selection for any given outlet is proportional to that outlet's share in total expenditures in the survey area for the item category in question. This is done so that the price quotes for selected items are obtained at outlets which are representative of the places that consumers made their purchases and also because the outlet is assumed to be an important characteristic of the purchase and component of price change. It follows from this assumption that differences in prices of the same item in different outlets must represent differences in aspects of the purchase such as quality of service or convenience of location and that consumers will pay the same proportional difference over time for these other aspects. When this assumption does not hold, such as when some outlets grow faster than others, the methodology will prevent adequate accounting in two ways: the current methodology will not adequately provide for obtaining more price quotes or give more importance to the more favored outlets, nor does it provide for direct comparison of the quality differences in purchasing the same item at two different places.

There is a process to "refresh" items and outlets sampled, called sample rotation, which generates a sample of specific items each of which had a probability of selection into the sample proportional to its share in recent consumer expenditures. Approximately 20 percent of the sample is rotated every year such that full rotation takes 5 years. The items rotated in are not directly compared to those they replaced. The procedure assumes that at the time of rotation, the original item and the one rotated in have the same quality adjusted price.

BLS procedure provides for selecting alternative items to be priced when the previously priced item is sold out, discontinued or otherwise permanently unavailable. The field agent is given guidelines to use in selecting the replacement or substitute item within the same ELI and a judgment is made as to the comparability of the specifications. (However, there is no provision to assure that the replacement is the product which has taken market share from the one that has disappeared.) <sup>14</sup> When the substitute is determined to be non-comparable, BLS most often assumes that the quality difference accounts for the price difference, net of the price change since the last pricing period for similar items. <sup>15</sup> In some cases, attempts are made to measure the quality differences. Notice that it is the disappearance of an item which triggers the mechanism to price a substitute.

Prices of new goods not falling within an established stratum, which are introduced after the base period and therefore not in the reference market basket, are not given special preference in item substitution and sample rotation, and consequently are often not included in the index until the subsequent decadal revision. <sup>16</sup> (Moreover, the impact of new goods is not measured retrospectively because the CPI is not revised historically. <sup>17</sup>) Frequently cited examples of important new products which were not introduced until many years after their introduction are air conditioners and VCRs. <sup>18</sup> Cellular telephones will be included in the 1998 revision of the CPI.

While the methodology does not ensure the introduction of new products until the market basket is updated, improving the timeliness through more frequent updates of the market basket solves only part of the problem. Direct comparisons of the quality of new products with those with which they compete is often difficult. Furthermore, proper accounting of the impact of new products often requires comparisons with products in other item groups. The current item structure prevents the CPI from fully capturing the effects of a drug replacing surgery, of electronic information services replacing newspapers, of automobile leasing competing with purchases, of video rentals replacing cinema attendance. Over time, price changes in successful products will be given greater weight in the CPI, but full measurement of the price impacts across item groups is not possible when close substitutes are in different item groups.

<sup>19</sup> — Although the item structure has several purposes, index estimation is the most important.

Prices for specific goods and services at specific outlets in specific locations are combined into item group-area indexes and these indexes are further aggregated by weighting them together either up through the item classification structure or by geographic area, to form a national CPI. The weights are derived from the Point-Of-Purchase Survey, the Consumer Expenditure Survey (which contains only modest detail) and from the statistical approach used in initiating specific commodities or services at the selected outlets. The design does not provide for collecting changes in quantities over time (since the market basket is assumed to be unchanged, this is not necessary to construct the CPI).

The use of arithmetic means to combine price changes within item groupings, for example different types of apples, implements the restriction that quantity weights do not change when prices change.<sup>20</sup> — The arithmetic mean fails certain common sense tests, as discussed in the next section.

The greater the substitutability of the items whose prices are combined this way, the greater the resulting substitution bias in the index. An alternative to assuming no change in quantities is to assume no change in expenditure shares. This can be accomplished through the use of geometric means, which effects a price increase that is proportionally offset by a quantity decrease. For example, if a ten percent increase in the price of granny smith apples were associated with a ten percent reduction in the quantity purchased, geometric means would be the appropriate way to capture the market response. If there were no quantity change associated with the price increase then arithmetic means would be appropriate. In the case of granny smith apples, the availability of other varieties of apples may yield a market response to quantities that more than offsets the price increase. When this happens, the use of geometric means understates the market response.

It is worth noting that the published geographic indexes do not provide comparisons of the price level across geographic areas; rather, they provide comparisons of rates of change in the CPI. Clearly, if the rates of change are different, then the levels must also differ at some point. Indeed, the differences in levels would be of significant interest as a comparison of the cost of living across geographic areas. Yet the methodology does not provide such comparisons. Geographic areas play an important role in the sampling design, however, geographic area indexes as they are constructed today serve no other purpose than a step in aggregation, en route to a national CPI.

In summary, sampling techniques are used to determine which items are priced at which outlets. The methodology requires allegiance to the concept of a fixed market basket which by design does not change item category weights until the market basket is updated, historically every ten years or so, and hence fails to capture some new products. Neither does it make direct comparisons of the purchase experience at different outlets, by assumption not capturing the lower prices to which consumers respond making some outlets grow faster than others. The most detailed level at which price indexes are constructed is for 207 item groups for each geographic area. Geographic area price indexes are constructed to provide estimates of price change in specific geographic areas en route to the national CPI, but they do not provide inter-area comparisons of the cost of living. Price indexes are successively combined into broader categories until a national CPI is reached.

In conclusion, improving the CPI as a measure of the cost of living requires addressing a range of issues beginning with revisiting critical assumptions, adjusting resource optimization criteria, and abandoning the Laspeyres index formula. The Commission's recommendations are presented in Section VIII.

#### **IV. The Consumer Price Index And A Cost Of Living Index: Measurement Issues**

A cost of living index is a comparison of the minimum expenditure required to achieve the same level of well-being (also known as welfare, utility, standard-of-living) across two different sets of prices. Most often it is thought of as a comparison between two points of time. As with any practical application of theory to index number production,

estimating a cost of living index requires assumptions, a methodology, data gathering processes and index number construction.

There are two sets of potential biases in the CPI: biases relative to an "ideal" cost of living index and biases which arise within its own terms of reference. The strength of the CPI is in the underlying simplicity of its concept: pricing a fixed (but representative) market basket of goods and services over time. Its weakness follows from the same conception: the "fixed basket" becomes less and less representative over time as consumers respond to price changes and new choices.

Consumers respond to price changes by substituting away from products that have become more expensive and toward goods whose prices have declined relatively. As the world changes, they are faced with new choices in shopping outlets, varieties, and entirely new goods and services, and respond to these as well. These changes make the previously "fixed basket" increasingly irrelevant.

In trying to keep true to its concept in a rapidly changing world, the current CPI procedures encounter difficulties. Biases result when they ignore some of these changes such as the appearance of discounters, and also when they try to do something about them such as when items are rotated out of the sample and replaced with new items. Attempting to capture the changes in a way that tries to mimic the pricing of a "fixed basket" within a rather patchwork framework just cannot be done without introducing other problems into the resulting index. These different biases overlap and have been discussed under a number of headings: substitution bias; formula bias; outlet substitution bias; quality change; and new product bias.

The "pure" substitution bias is the easiest to illustrate. Consider a very stylized example, where we would like to compare an initial "base" period 1 and a subsequent period 2. For simplicity, consider a hypothetical situation where there are only two commodities: beef and chicken. In period 1, the prices per pound of beef and chicken are equal, at \$1, and so are the quantities consumed, at 1 lb. Total expenditure is therefore \$2. In period 2, beef is twice as expensive as chicken (\$1.60 vs. \$0.80 per

pound), and much more chicken (2 lb.) than beef (0.8 lb.) is consumed, as the consumer substitutes the relatively less expensive chicken for beef. Total expenditure in period 2 is \$2.88. The relevant data are presented in Table 1. How can we compare the two situations? Actually, there are several methods, each asking slightly different questions and therefore, not surprisingly, giving different answers.<sup>21</sup>

<b>TABLE 1: HYPOTHETICAL EXAMPLE OF SUBSTITUTION BIAS</b>								
	Price in Period 1	Quantity in Period 1	Price in Period 2	Quantity in Period 2	Price Relatives		Relative Weights	
					P2/P1	P1/P2	1	2
Beef	1	1	1.6	0.8	1.6	0.63	0.5	0.43
Chicken	1	1	0.8	2.0	0.8	1.25	0.5	0.57

The simplest comparison is to ask "How much more must I spend in my current situation (period 2) to purchase the same quantities that I purchased initially (in period 1)?"<sup>22</sup> This is the question asked by the CPI. The price index for period 2 relative to period 1 uses the initial period 1 basket of consumption as the weights in the computation. To buy 1 lb. of beef and 1 lb. of chicken in period 2 costs \$2.40. The price index for period 2 relative to period 1 is 1.20 (2.40/2.00), that is a 20 % increase. Intuitively, it is easy to understand why such a computation imparts an upward (substitution) bias to the measure of the change in the true cost of living. It assumes the consumer does not substitute (cheaper) chicken for beef. In the real world, as in the hypothetical example, consumers change their spending patterns in response to changes in relative prices and, hence, partially insulate themselves from price movements.

An alternative approach would be to ask the question "How much more am I spending in my current situation (period 2) than I would have spent for the same goods and services at the prices that prevailed initially (in period 1)?"<sup>23</sup> This price index compares expenditures in period 2 (\$2.88) with what it would cost to buy the current (period 2) market basket at the initial prices (\$0.80 for the beef plus \$2.00 for the chicken equals \$2.80). This price index is 1.03, that is only a 3% increase. This approach understates the rise in the true cost of living as it overstates substitution.

The idea underlying a cost of living index is to allow for the substitution that follows relative price changes. The question answered by a cost of living index is: "How much would we need to increase (or decrease) initial (period 1) expenditure in order to make the consumer as well off as in the subsequent period (period 2)." Although the answer to this question might appear to require detailed knowledge of a consumer's preferences, an excellent approximation can be obtained by using a "superlative" index formula instead of the traditional fixed weight index employed in the CPI.

The concept of a superlative index number was introduced by the American economist, Irving Fisher (1922), to describe index numbers that met certain reasonable criteria and thus agreed closely with his "ideal" index, described below.<sup>24</sup> This concept was generalized by the Canadian economist, Erwin Diewert (1976), and used to describe any index number formula that provides a satisfactory approximation to an underlying economic index, such as a cost of living index.<sup>25</sup> The CPI is based on a fixed weight index formula that does not provide such an approximation, fails to meet these sensible criteria and worse yet is known to be biased upward. A superlative index requires the same information on prices and quantities as a fixed weight index, but involves interpolating between the two periods rather than treating one of them as the "base" period. There are two ways of doing this.

The first approach to interpolating between time periods is to use the geometric mean of the two fixed weight indexes -- using the initial period and the subsequent period as "base" periods. The geometric mean is the square root of the product of the two indexes. This is the ideal index originated by Irving Fisher (1922) and now called the "Fisher ideal index" in his honor. In our example, this comes to 1.11, an 11% increase. By comparison the CPI-type fixed weight index, treating period 1 as the "base" period, is biased upward by 9% (1.20 minus 1.11). Alternatively, a fixed weight index with period 2 as the "base" period is biased downward by 8% (1.03 minus 1.11). The Fisher ideal index is employed by the Bureau of Economic Analysis in compiling data on the U.S. national income and product accounts.

An alternative approach to interpolation is to use a weighted average of the growth

rates in prices with relative weights equal to the average of the weights in the two periods. This is called the "Tornqvist" index in honor of one of its originators, the Finnish statistician Leo Tornqvist (1936).<sup>40</sup> In our example, this is 1.10, a 10% increase. We conclude that the two superlative index formulas yield very similar approximations to the cost of living index. Estimates of the biases of the two fixed weight indexes are also similar. The BLS has compared a fixed weight index with the Fisher ideal and Tornqvist indexes to assess the bias in the fixed weight index as a measure of changes in the cost of living.

How large are substitution biases in the CPI? To answer this question we must take into account the hierarchical nature of the construction of the CPI described above. It is useful to focus initially on Upper Level Substitution Bias, which occurs when indexes for the 207 item groups and 44 areas are aggregated to form the CPI. The BLS uses a fixed weight index for this purpose (with weights derived from the Consumer Expenditure Survey (CES), a survey of household expenditure patterns), and hence ignores substitutions of chicken for beef, apples for oranges, etc. The BLS has measured this form of substitution bias by comparing a fixed weight index with an index generated by one of the interpolation methods we have described. Estimates are presented in Section VI.

The second type of substitution bias in the CPI is Lower Level Substitution Bias, which occurs when prices for the approximately 71,000 goods and services and information on housing costs are used to form indexes for the 207 items and 44 areas. This part of the index construction involves probability sampling with probabilities derived from the CES and the Point-of-Purchase Survey (POPS) of retail establishments in order to reflect the likelihood of purchases of individual items at specific retail outlets. It is useful to think of this as an alternative fixed weight index with probabilities playing the role of expenditure weights. Since 1978 the formula at the lower level of index construction has been closely analogous to the fixed weight index at the upper level. The use of arithmetic means to aggregate the price changes assumes no substitution to changes in the relative prices of the specific commodities or services within the lowest grouping, e.g., when the price of granny smith apples rises,

it assumes no substitution, of say, delicious apples. The BLS has measured Lower Level Substitution Bias by comparing this fixed weight index with a geometrically weighted average of prices at the lower level.

A major difficulty with a fixed weight index at the lower level of index construction is the failure of time reversibility. This simple and intuitive requirement or "test" for an index number is that the index should remain the same if the underlying prices undergo a reversal. For example, suppose that the price of beef in Table 1 rises from 1.0 in Period 1 to 1.6 in Period 2, but then falls back to 1.0 in Period 3, reversing the change that took place between Periods 1 and 2. A fixed weight index increases by 60% between periods 1 and 2, but decreases by only 37.5% between periods 1 and 3, so that the increase in the "beef" index between periods 1 and 3 is 22.5% or 11.25% per period, rather than zero, as required for time reversal.

A geometric average satisfies the time reversal test, since it is based on the square root of the product of the price ratios between periods. In the example of the beef price from Table 1, the price ratio between Period 1 and Period 2 is 1.6, while the price ratio between Period 2 and 3 is .625. The product of these two price ratios is one, as required for time reversal, so that the average price increase is zero per period. The time reversal property has led to widespread use of of geometric averages as a standard for comparison of different approaches at the lower level of index number construction. For example, Moulton and Smedley (1995) have compared the BLS fixed weight approach at this level with a weighted geometric approach.<sup>27</sup>

Diewert (1995) has provided a detailed review of the properties of alternative approaches to index number construction at the lower level.<sup>28</sup> These include time reversal, as well as other reasonable requirements for index numbers at the lower level. Shapiro and Wilcox (1996b) have provided an elegant rationale for the geometric approach based on the correlation of relative prices over time.<sup>29</sup> Provided that this correlation is small, a modification of the geometric mean is approximately unbiased for the underlying cost of living index, and this characterization does not require information about the underlying system of consumer's preferences.

Modified geometric means have been widely used as a standard for evaluating methods for index number construction at the lower level. Diewert (1995) gives a useful review of the empirical studies. In addition to the work of Moulton and Smedley (1995), Carruthers, Sellwood, and Ward (1980) have conducted a study of this type for the U.K., Schultz (1994) for Canada, Dalen (1994) for Sweden, and Woolford (1994) for Australia.<sup>30</sup> These studies show that fixed weight indexes, like those used by BLS, are biased upward; the order of magnitude of the bias is similar to that suggested by the study of Moulton and Smedley (1995) for the U.S. These problems have led an increasing number of statistical agencies, such as Statistics Canada, to follow Irving Fisher's (1922) advice and jettison the arithmetic mean in favor of the geometric mean.

A relatively subtle problem developed in implementing the fixed weight index at the lower level. When sample items are replaced by substitute items for which no previous price observations are available, base period prices for the substitute items must be "imputed" to fill this gap. The procedure adopted by BLS for doing this had the effect of linking the weights for the substitute items to the prices used in the CPI and produced a bias that is an important component of Lower Level Substitution Bias. This problem also arises during rotations of items included in the sample of 70,000 prices for goods and services and the sample of housing costs.

An estimate of the overall Lower Level Substitution Bias is give by the difference between the fixed weight index and a geometrically weighted average, where the fixed weight index is based on the methods for price imputation introduced by BLS in 1978. In 1995 and 1996 BLS introduced new procedures based on "seasoning" the price estimates. Seasoning involves lengthening the period between a price imputation and the period when an item is actually introduced into the CPI. By lengthening this period the link between weights and prices for individual items can be broken and the bias reduced. However, the bias associated with the fixed weight formula remains.<sup>31</sup>

Our Interim Report anticipated that what we called "formula bias" and now refer to as Lower Level Substitution Bias would be eliminated by BLS. The BLS did alter its

procedures by introducing "seasoning" where appropriate; while this eliminated bias due to methods for price imputation, it did not affect the bias due to the use of a fixed weight formula at the lower level. Accordingly, we have recommended below that BLS should replace the fixed weight formula by a geometrically weighted average. This has been tested and found to be feasible in an important study by Moulton and Smedley (1995).<sup>52</sup> Additional work is currently underway to extend the period of this study.

The introduction by BLS of a fixed weight index at the lower level of aggregation was viewed at the time as introducing consistency of indexing at both upper and lower levels of aggregation. However, the disadvantages of the fixed weight approach at the upper level carry over to the lower level. A superlative index formula is required to provide a satisfactory approximation to the underlying cost of living index at the upper level. This avoids the bias associated with the fixed weight index formula employed in the CPI. Similarly, lacking quantity or expenditure information at the lower level, a good approximation to the underlying cost of living index is obtained from a geometrically weighted average formula.

Just as consumers change the goods they purchase in response to changes in relative prices, as in the beef and chicken example, so do they change the location where they make their purchases. The opening of a new discount store outlet may give consumers the opportunity to purchase at a lower price than before. At present, the CPI procedures ignore such reductions that occur when consumers change outlets. However, if consumers cared only about obtaining goods at the lowest price, then we would observe all goods sold at the same price at all outlets. Instead, we observe low prices at discount stores and warehouse clubs at the same time as medium prices at supermarkets and higher prices at convenience stores. Evidently, consumers care not only about prices, but the level of services such as availability of clerks, wrapping services, and the distance between home and alternative outlets.

Current procedures in the CPI ignore price changes when consumers switch outlets. This incorporates into the CPI the implicit assumption that price differentials among

outlets entirely reflect the differences in service quality. This approach would be legitimate if the economy stood still with a stable set of outlets providing alternative levels of service quality. However, there has been a continuous increase in the market share of discount stores as more efficient technologies of distribution allow low price outlets to expand while older, higher priced outlets have contracted and in some cases gone out of business. This shift in market share indicated that many consumers respond to price differentials and do not consider them to be fully offset by differences in service quality. Completely ignoring all differences in service quality by incorporating all such price reductions into the CPI would err in the opposite direction. Further research is required to disentangle true changes in prices from changes in service quality. This problem is analogous to the need to disentangle the changes in prices from changes in product quality.

Quality change and new goods present the most difficult problems for measurement. They include capturing the introduction of new products in a timely manner; making direct quality comparisons of new products with existing ones; making direct quality comparisons of new products with other products against which they compete (in other classification groupings such as a new drug and the surgical treatment it replaces); and capturing the combined impact of quality and substitution as these new products displace others within and across their classification grouping.

A well-known expert on price indexes has stated the general issue clearly: "...heterogeneity in economics pertains to transactions, and not just the physical description of the product. Whenever two transactions involve different bundles of explicit or implicit attributes, they differ qualitatively. Differences in terms of sale, services provided with the sale,...are exactly identical from the economics of the matter, to physical changes that we normally think of as "quality change" (Triplett (1990)).

For example, it is not just what is purchased where (and how), but possibly also when that matters. There may also be a time of week bias. The BLS does not collect prices on weekends and holidays when certain items and types of outlets disproportionately

run sales.<sup>53</sup> There appears to have been a sizeable increase in the fraction of purchases made on weekends and holidays perhaps reflecting the increased prevalence of two earner families. We know of no systematic study of this issue and urge the BLS to conduct the research necessary to examine it thoroughly, perhaps with scanner data.

A full treatment of these issues reinforces the problem of focusing on the "average" or "representative" consumer. Different consumers have different tastes and time costs, and hence value the appearance of new outlets and new products differentially, with some (the majority) becoming better off with supermarkets and others losing out as the corner grocery store disappears. The CPI is not equipped to account for special characteristics of different consumers or groups of consumers.<sup>54 55</sup>

There are still other issues that would in principle apply to obtaining a true cost of living index (COLI). Consider two examples: the negative effects of higher crime rates and the concomitant purchases of security devices and higher insurance premiums and the positive effects of improvements in information technology that permit a parent to work at home when a child is ill. Surely these would enter a calculation of "the minimum expenditure necessary to be at least as well off." Section VII below explores some of these problems.

## **V. Quality Change And New Products**

### ***Introduction***

The difficult questions posed by quality change and the continuing arrival of new products have been called the "house-to-house combat of price measurement." In this section we will treat new product bias as a component of quality change bias and will not attempt to break down our overall bias estimate into the separate contributions of quality change bias and new product bias.

Quality changes have occurred at a rapid rate for some products but not others. The CPI has done a better job capturing the effect of quality change for some products than others. The CPI has introduced some new products faster than others. Because the magnitude of quality change bias differs so much across product categories, any overall evaluation of the magnitude of quality change bias must be conducted "down in the trenches", taking individual categories of consumer expenditure, assessing quality change bias for each category, and then aggregating using appropriate weights. Further complicating the analysis is that quality change bias, assessed at the level of individual products, appears to have changed significantly over time. For instance, important improvements in BLS methodology largely or entirely eliminated an upward bias in the CPI for new automobiles prior to the mid-1960s and a downward bias for apparel after the mid-1980s. Likewise, an important source of downward bias in the CPI rent index was eliminated in the late 1980s.<sup>30</sup>

Previous evaluations of quality change bias, e.g., Shapiro and Wilcox (1996c) and Lebow, Roberts, and Stockton (1994) have tended to take bias estimates from earlier research on particular products, e.g., consumer appliances or automobiles, apply that bias estimate with the weight of those products in the CPI, and assumes that in the rest of the CPI the rate of quality change is zero. We do not view that approach as likely to emerge with a neutral evaluation of the bias, simply because the evaluation that the rest of the CPI is unbiased represents an extreme one-sided answer to the question as to whether the components of the CPI subject to relatively little research are biased. They may be as likely to be subject to the average rate of bias of those components which have been subject to careful research as to no bias at all. In this section we evaluate the CPI component-by-component and extrapolate research on bias from one category to another when the categories seem related. Nevertheless, we attribute bias estimates of zero to a number of categories which seem quite dissimilar to those categories subject to intensive research, or where unmeasured quality change and new products have been relatively unimportant.

While the problem of bias due to quality change and new products can be largely separated from the other forms of bias considered above, this is not entirely possible.

Evidence on quality change bias developed in other studies, for instance Gordon (1990), is based on an attempt to measure prices directly from sources independent of BLS price quotations, using such sources as mail order catalogues and *Consumer Reports*.<sup>37</sup> However, any differences between these independent indexes and the CPI for the same goods may reflect not just quality change and new product bias, but also traditional substitution bias (since the mix of products and models shifts faster in the alternative source than the CPI), outlet substitution bias (since alternative price quotes are often an average of market prices which adjusts for the changing mix of discount stores), and formula bias (since the alternative indexes are free from the formula bias problems discussed previously).<sup>38</sup>

### ***Conceptual Issues***

The difficulty created by quality change in existing products, and by the introduction of new products, is highlighted by returning to the definition of a cost of living index: a comparison in two time periods of the minimum expenditure required to achieve the same level of well-being. What does the "same level" mean when the models of a given product available in the second time period embody different quality attributes than in the first time period? And, an even more profound difficulty, what does the "same level" mean when entirely new products are introduced that were unavailable in the first time period?

A pervasive phenomenon called the "product cycle" is critical in assessing the issue of new product bias in the CPI and applies as well to new models of existing products. A typical new product is introduced at a relatively high price with sales at a low volume. Soon improvements in manufacturing techniques and increasing sales allow prices to be reduced and quality to be improved. For instance, the VCR was introduced in the late 1970s at a price of \$1,000 with clumsy electromechanical controls; by the mid 1980s the price had fallen to \$200 and controls were electronic, with extensive preprogramming capabilities. Later on in the product cycle, the product will mature and eventually will increase in price more rapidly than the average product of its class. The sequence is easily visualized as a "U"-shaped curve -- the price of any given

product relative to the consumer market basket starts high, then goes down, is flat for a while, and then goes back up. To the extent that the CPI overweights mature products and underweights new products, it will tend to have an upward bias. Some recent academic research, notably Berndt, Griliches, and Rosett (1993) and Berndt, Cockburn, and Griliches (1996), computes alternative price indexes with the mix of prescription drugs actually sold and the limited and older sample contained in the CPI, and this research attributes a significant upward bias to the CPI on the grounds of its lateness in introducing the mix of models and varieties actually sold.

An important criterion for the assessment of quality change and new product bias is the evolution of market shares for particular models and products. When a new model is introduced that is more expensive than an old model, but it gains market share, we can conclude that it was superior in quality to the old model by more than the differential in price between the two.

The same criterion helps us deal with outlet substitution bias. When consumers shift from traditional supermarkets to new, more expensive specialized food markets offering an improved selection or variety of produce, we can deduce that consumers are better off. The fact that Wal-Mart both charges lower prices and has become by far the largest retail chain over the past 15 years indicates that consumers do not consider the lower Wal-Mart prices to be offset by inferior service, as implicitly assumed by the CPI, but rather that consumers view Wal-Mart to offer a superior combination of prices and service to the previously available mix of outlets. The fact that convenience stores like 7-11 both charge higher prices and have gained market share indicates that consumers view convenience stores as providing a value of extra convenience that is worth more than the extra price that they charge. Many consumers shop at both Wal-Mart and convenience stores, paying both lower and higher prices on particular items than with the previous mix of stores, and the shift in market share suggests that the new mix is an improvement. The same evaluation can be made of restaurants, where consumers have shifted toward low-priced fast food outlets like McDonalds, medium-priced franchises like Olive Garden and Red Lobster, and in some urban areas, sophisticated high-priced restaurants specializing in Tuscan, Thai, and other

ethnic food specialties. An important strand of academic research on such diverse products as medical imaging devices (Trajtenberg, 1990) and breakfast cereal (Hausman, 1996) attributes substantial value to increases in product variety. Thus, the "value of variety" is critically important in our assessment both of outlet substitution bias and, in this section, of quality change and new product bias.

### ***BLS Methodology***

Our discussion of quality change and new product bias begins with a review of the methods used by the CPI to handle quality changes in existing products and then turns to problems posed by new products. The BLS has five different methods to cope with a model change for an existing product.

The "direct comparison" method treats all of the observed price change between the old model and the new model as a change in price and none as a change in quality. There is no necessary bias, because quality can decrease as well as increase. But in practice most goods tend to undergo steady improvement, and often a better model is introduced with no change in price, causing the quality change to be missed entirely.

The "deletion" method makes no comparison at all between the prices of the old and new model. Instead, the weight attributable to this product is applied to the average price change of other products in the same commodity classification. To the extent that the deletion method is used, the CPI consists disproportionately of commodities of constant quality which may be further along in the product cycle.

The "linking" method can be used if the new and old model are sold simultaneously. In this case the price differential between the two models at the time of introduction of the new model can be used as an estimate of the value of the quality differential between the two models. As indicated above, this can lead to an understatement of quality change if the new model gains market share. Also, a quality improvement in the new model can occur even if it costs less or the same as the old model, as in the case of the VCR where the price fell continuously while programming capability and reproduction quality improved.

The "cost estimation" method attempts to establish the cost of the extra attributes of the new model. Problems in practice with the costing method have been its infrequency of use, and the fact that it has been applied disproportionately in the case of automobiles relative to other products. This raises the possibility that there is a spurious upward "drift" in the relative price of other products relative to automobiles due to an uneven application of the costing method. An emerging source of upward bias is that products like automobiles are benefitting from the improved quality of materials like steel (which does not rust as it once did) and tires (which last many more miles). To the extent that some of these inputs to the auto production process are experiencing quality improvements of their own in excess of differences in cost, these will not be picked up by the BLS cost-based quality estimation procedure.

Thus far, the CPI has introduced only in its apparel category an alternative methodology called the "hedonic regression method" for estimating the value of quality change. The hedonic approach can be viewed as an alternative method to manufacturers' cost estimates in making quality change adjustments. It assumes that the price of a product observed at a given time is a function of its quality characteristics, and it estimates the imputed prices of such characteristics by regressing the prices of different models of the product on their differing embodied quantities of characteristics. Thus the hedonic approach is less a new method than an alternative to cost estimates to be used when practical factors make it more suitable than the conventional method.

By their very nature hedonic indexes require large amounts of data. Given the thousands of separate products that are produced in any modern industrial society, the need to collect a full cross-section of data on each product presents a substantial obstacle to the full-blown adoption of the hedonic technique. But in many cases the data already collected by CPI field agents can be used for hedonic regression analysis; this is already done in the case of apparel.

Another possible objection is that it is impossible to construct a hedonic index in the timely fashion required by the CPI, with its orientation to producing within a few

weeks an estimate of month-to-month price changes that can never be revised. But this ignores the fact that coefficients can be estimated on the basis of historical data, and these previously estimated coefficients can be used to evaluate quality change when a new model is introduced. This approach would be particularly suitable for product categories subject to a rapid succession of new model introductions, notably TV sets and personal computers.

This list of BLS methods reveals at least four potential sources of upward bias: the use of the direct comparison method that does not address the quality issue at all, the use of the deletion method that bases price change on models that are unchanged in quality and may be further along in the product cycle, the use of the linking method when quality improvements are greater than the price differential across models, and the use of the cost method which may miss quality improvements achieved by those firms which supply better materials and inputs to producers of final goods.

A potentially greater difficulty is that the CPI makes no attempt to create systematic estimates of the value of quality improvements which increase consumer welfare without raising the price of products. For instance, many consumer electronic products and household appliances have experienced a reduction in the incidence of repairs and in electricity use, and few if any of these improvements have been taken into account by the CPI. The increased longevity of automobiles (cited below), appliances, and other products introduces a similar source of bias.

### ***New Product Bias***

We turn now to the issue of new product bias. There is no debate regarding the reality of the product cycle, and nobody debates the fact that the CPI introduces many products late, thus missing much of the price decline that typically happens in the first phase of the product cycle. An extreme example involves room air conditioners, which were widely sold in 1951, but not introduced into the CPI until 1964, 13 years later. More recently, the microwave oven was introduced into the CPI in 1978 and the VCR and personal computer in 1987, years after they were first sold in the marketplace. As

an even more contemporaneous example, there are currently 36 million cellular phones in use in the United States, but as yet the CPI has no price index for cellular phones. Thus none of the benefit to consumers of being able to keep track more easily of children, spouses, or of aged parents has yet received any credit in our national measures of inflation, real output, or productivity. Even more recently, there are more than 40 million cellular phone subscribers in the U.S., but the cellular phone has yet to be introduced into the CPI.<sup>59</sup>

A second aspect of new product bias results from a narrow definition of a commodity. When a new product is finally introduced into the CPI, no comparison is made of the price and quality of the new product with the price and quality of an old product that performed the same function. For instance, people flock to rent videos, but the declining price of seeing a movie at home, as compared to going out to a theater, is not taken into account in the CPI. Similarly, the CPI missed the replacement of electric typewriters by electronic typewriters and then PCs with word-processing and spell-checking capability, or CD-ROM encyclopedias that cost far less than old-fashioned bound-book versions and eliminate many trips to the library. Inevitably, however, many new products embody genuinely new characteristics that have no previous counterpart. Electronic mail that provides a new set of bonds and communication between parents and their children who are off at college and cellular telephones that make possible virtually continuous contact with a sick child or aged parent are but two examples.

This discussion of new products leads inevitably to deeper questions about changes in the standard of living of the average American. Positive changes made possible by consumer electronics need to be weighed against increasing crime, pollution, and other "bads." We return to these issues in Section VII below.

### ***Quality Change and New Product Bias by Product Category***

Because quality change bias differs in magnitude, direction, and timing across product categories, the only way to narrow the range of uncertainty of the magnitude of quality

change bias is to examine the available evidence, category by category. Table 2 is designed to provide a guide to this assessment. The left-hand column lists each major product category within the CPI next to its "relative importance", i.e., percentage weight, in December, 1995. In this section we review the available evidence on bias related to quality change and new products, by category.

In some categories there is little if any published evidence that allows us to reach a determination. However, we do not follow previous research by assuming that in these categories the overall bias due to quality change and new products is necessarily zero. Instead, we discuss the likely direction of bias in the context of the definition of a cost of living index: a comparison in two time periods of the minimum expenditure required to achieve the same level of well-being.

**1. Food and beverages.** The most dramatic evidence of upward bias in the food and beverages category was produced by Reinsdorf (1993), who found during the 1980-90 period an annual rate of change of average price paid for 50 narrowly defined commodities that was fully 2.0 percent per annum slower than the CPI for the same product categories. While Reinsdorf thought at the time that this difference reflected outlet substitution bias, in fact he later concluded that the difference represented a mix of formula bias and outlet substitution bias. Whatever the interpretation of Reinsdorf's study, it does not represent evidence on quality change, since his commodities were chosen to be identical to those priced in the CPI.

Besides his study, there is little if any published evidence on the food category, other than Hausman's (1996) attempt to establish the value for the introduction of a new variety of breakfast cereal. Perhaps more important than new varieties of packaged goods has been a wave of technological improvements that has greatly increased the variety of fresh fruits and vegetables available in the typical supermarket during the winter months, and a trend toward more services provided in supermarkets, eliminating the need to travel to small specialty shops, especially fresh fish markets and deli counters preparing fresh-cooked food. How much would a consumer pay to have the privilege of choosing from the variety of items available in today's

supermarket instead of being constrained to the much more limited variety available 30 years ago? A conservative estimate of the value of extra variety and convenience might be 10 percent for food consumed at home other than produce, 20 percent for produce where the increased variety in winter (as well as summer farmers' markets) has been so notable, and 5 percent for alcoholic beverages where imported beer, microbreweries, and a greatly improved distribution of imported wines from all over the world have improved the standard of living. Increased variety and convenience in food away from home, in every price category from McDonalds to luxury restaurants (as discussed above), can also be credited with a 10 percent premium. The annual rates of bias in Table 2 are calculated by converting these assumed premia to annual geometric growth rates over the past 30 years.

2. **Housing.** By far the largest single weight in the CPI is given to housing component, and within that to shelter. The shelter component shifted to a rental equivalence approach in 1983, and the CPI-U-X1 index represents an attempt to provide a consistent treatment of housing using the rental equivalence concept back to 1967. The annual rate of change of the CPI shelter index exceeds that of the CPI residential rent index by 2.33 percent per annum from 1967 to 1983, and correspondingly the annual rate of change of the official CPI-U exceeds that of CPI-U-X1 by 0.52 percent per annum over the same interval.<sup>40</sup> The BLS has also shifted methodology in 1995 to correct formula bias and in 1988 to correct an "aging bias" that resulted from pricing in successive periods housing units that were becoming progressively older. Randolph (1988) estimates this pre-1988 aging bias at 0.3 percent per annum, a concept that represents the effect of depreciation net of any maintenance and renovation expenditures.

First, we register our skepticism that the Randolph aging bias should be considered a bias in its entirety. Older units rent for less than new units for two reasons. First, they may physically deteriorate by more than is offset by repairs and maintenance. But, second, they may lose value as newer units come on the market containing amenities such as central air conditioning. Such economic obsolescence does not represent a decline in the quality of the service provided by the older apartments, but rather

represents the result of the fact that the income elasticity of demand for shelter amenities is positive, and people expect higher quality in apartments and houses as the nation's per capita income increases. An exact analogy is the introduction of the jet plane, discussed in detail by Gordon (1990). The quality of the ride on a propellor-driven DC-7 did not decline when the pure-jet DC-8 was introduced in 1958. Rather, consumers valued the ride on the jet plane so highly that the demand for flights on the DC-7 vanished. The DC-7 was scrapped prematurely, within five to ten years after the introduction of the jets. Consumers gained the entire surplus from the transition from propellor to jet planes for long-distance air travel, and the declining rents of older apartments represent a less dramatic example of the same phenomenon. Thus far there has been little investigation into quality change in the apartments included in the CPI rent survey. The "CPI methods hold most housing quality constant by measuring rent changes longitudinally for a cross-section of housing units" (Randolph, 1988, p. 359). That is, rent changes on a given unit are followed through time, and alternative units are rotated in, with the overlap handled by deletion. If there is a general tendency for more recently constructed units to have more and better appliances, central air conditioning, and other amenities that were not present in previous decades, there is the possibility of an upward bias in the CPI rental index if consumers value these amenities at more than their extra cost. The continuous movement of households to newer apartment complexes in suburbs and in the Sunbelt may be part of a process by which housing quality steadily improves. The "market share" test suggests that many households prefer new sunbelt apartments to older types of apartment in central cities in the north central and northeastern states.

The U. S. Census Current Housing Reports report median monthly rent of all rental occupied units. The ratio for 1993 to 1976 is 2.92 (\$487/\$167). The CPI rental index ratio (not adjusted for formula or aging bias) for the same years is 2.46. The implied annual difference in growth rates for the CPI is -1.00 percent per year. An alternative comparison for 1973-88 yields a difference of -1.10 percent per year.<sup>41</sup>

While only limited data are available on the quality of rental units, there is evidence that rental units have improved in quality at approximately the pace of owner-occupied

units, for which more data are available. Two key measures have persuaded us of the comparability of rental and owner units (the CPI uses rent indexes for both the rental and owner-occupied segments of housing, so these findings support the CPI choice). First, between 1970 and 1993 the mean number of rooms increased by 9.7 percent in all occupied units (of which about 1/3 were rental units), while the mean number of rooms in rental units increased by a similar 7.8 percent. Perhaps more important, the number of rooms per person increased by 30.2 percent in all units and 27.0 percent for rental units.<sup>42</sup> This set of comparisons supports the view that quality has improved at approximately the same rate in rental and owner-occupied units, and that we can use some of the available data on the totality of occupied units to reach a judgment on the extent of quality change.

While the best data are available for newly constructed units, some important data are available for the entire stock of existing units. For the entire stock of existing rental units alone, the mean number of bathrooms increased by 23.3 percent between 1970 and 1993. And for the entire stock of all units, the fraction containing central air conditioning increased from 10.8 to 41.7 percent.

Further indication of the change in quality standards is indicated by changing characteristics of new single-family houses completed in 1993 compared to 1976: median square feet increased by 30 percent, bathrooms from 2.0 to 2.4, percentage with central air conditioning from 49 to 78, percentage with one or more fireplaces from 45 to 63, and percentage with a garage from 72 to 84.<sup>43</sup>

We have already determined that between 1976 and 1993 the average rent paid in the U.S. increased 1.0 percent faster than the CPI rent increase. To conclude that the CPI is unbiased, we would have to determine that the quality of the average rental unit increased by 1.0 percent per year over that period, or 18 percent over the entire period. From the evidence we have examined, we believe that 20 percent is a low-end estimate of the increase in the average size of apartments, which would support the conclusion that the average rent per square foot has increased no faster than the CPI. But also, we find convincing evidence that the average quality of apartments per square foot has

increased as well. The transition to central air conditioning proceeded at a rapid rate during the past two decades. Other amenities were added which increased the average quality of apartments, particularly swimming pools, health clubs, on-site free parking, and climate (since the mix of apartments shifted toward southern climates which reduced the impact of winter weather on tenants, particularly older tenants).

For the period since 1970 we find it plausible that the CPI accurately measures rent per square foot of apartment space, but its measure of shelter rent is upward biased by neglecting the increase in the quality of apartments per square foot. It is entirely natural that an increase in per-capita income would spill over into increased quality of housing, because there is no reason why housing size and quality should have an income elasticity of zero. The improved quality of appliances documented by Gordon (1990) applies to the shelter sector, since most apartments are now provided with relatively recent refrigerators, stoves or oven/cooktop combinations, dishwashers, and garbage disposals. The rental equivalent of these appliances must be substantial and they have been included in both new and older apartments mainly since 1955-60. A conservative estimate is that the total increase in apartment quality per square foot, including the rental value of all appliances, central air conditioning, and improved bathroom plumbing, and other amenities, amounts to 10 percent over the past 40 years, or 0.25 percent a year. Accordingly, Table 2 records an upward bias in the CPI of 0.25 percent per year for the shelter component, and this well may be an understatement.

For years before 1973, there is some evidence that the CPI rent index may be biased downward by more than can be explained by changes in quality. For instance, average annual rental expenditures for working class families in the CES increased from \$444 in 1950 to \$1803 in 1973, a ratio of 4.06, while the equivalent ratio for the unadjusted CPI rent index is only 1.93. This translates into a slower annual growth rate of the CPI of -3.24 percent per year. The same comparison for 1918 to 1950 yields an annual difference of -2.82 percent per year.<sup>44</sup> Without a measure of annual quality change per year, we cannot make a judgement on the magnitude of the bias, but the possibility that the CPI rental index incorporates a substantial downward bias prior to 1973 may help to explain the "Nordhaus thought experiment problem" identified above, namely

that backward extrapolation of substantial CPI bias for a century or more yields implausibly low levels of the standard of living during the 19th century. Further judgment on this issue must await the development of quantitative measures of the change in apartment quality between 1918 and 1973, although we note that there has obviously been a major improvement in quality since 1918, when only 36 percent of apartments had bathrooms and only 61 percent had inside water closets (Brown, 1994, Table 3.6A).

Turning now to other components of housing expenditure, there is no reason to suppose that the CPI has measured the price of fuel or electricity inaccurately, since these commodities are homogeneous and among the easiest to measure of any goods or services. However, when we think of why people prefer to live in the modern age and would (in most cases) not willingly choose to go back to the conditions of 70 years ago, the change in the nature of household heating fuel surely enters the calculation. In 1918, 80 percent of American homes were heated with coal and wood, which had to be stored and carried, and produced a fire that had to be tended, used a stove that had to be cleaned, and smoke that polluted the air.<sup>43</sup> Because the transition from coal and wood heat to other sources of fuel had been largely completed by the early 1970s, we do not include this major improvement in the quality of life as a source of recent bias in the CPI.

The rest of the weight in the CPI on housing is applied to a myriad of expenditures, each having a relatively small weight, including telephone service, refuse collection, cable TV, curtains, furniture, bedding, video and audio products, major household appliances, and a large number of miscellaneous items. Most of the CPI weight on "other utilities" is applied to local and long distance telephone service and cable television. Even if the CPI correctly tracked the prices of each of these items, quality change would be missed. There has been continuous improvement in the quality of telephone service (e.g., reduction of static and improvement in clarity), improved convenience (credit card pay phones, itemized billing), and a great increase in picture quality and consumer choice achieved by cable television viewed as a new product. The fact that more than 60 percent of American households are now wired for cable

TV, despite substantial monthly program fees, suggest that the development of cable TV has created a product yielding substantial consumer surplus. We conservatively estimate the quality bias connected with this category as 10 percent per decade, or 1.00 percent per year.<sup>46</sup>\_\_

The appliance and radio-TV category has been subject to more extensive research than any other category of consumer spending. Over the full period 1947-83 Gordon's detailed study (1990, p. 552), based on model-by-model comparisons from Consumer Reports, found an upward bias in the PCE deflator (which in turn is based on the CPI) of 3.22 percent per year for appliances and 5.94 percent for radio-TV. For the 1973-83 subperiod, the respective rates are 2.83 percent and 4.69 percent. These rates are applied in the CPI to a remarkably small fraction of consumption, just 0.8 percent according to Table 2. Consumer electronics alone, i.e., excluding electric appliances, recorded annual factory sales (i.e., net of retail markups) of \$55.9 billion in 1994, which amounted to 1.25 percent of nominal personal consumption expenditures.<sup>47</sup>\_\_ The 1995 share in PCE of final sales to consumers of audio and video equipment, including TV sets and VCRs, was also 1.25 percent, appliances contributed an additional 0.55 percent and personal computers an additional 0.33 percent, for a total weight in PCE of 2.13 percent, well over double the weight of the same products in the CPI.<sup>48</sup>\_\_

This small slice of personal consumption is the source of the largest annual rate of bias, with the possible exception of medical care. Our overall estimate of bias, based on Gordon's research, incorporates both quality-change bias and also new product bias, since his estimates of the overall bias take account of the fact that the quality-adjusted price of the VCR was declining at 30-40 percent per year in the early 1980s, prior to the introduction of this product into the CPI in 1987. Similarly, in recent years the price of personal computers purchased by consumers has been declining by at least 25 percent per year, but this has no impact at all, because home purchase of PCs were negligible in the CPI base period of 1982-84.

Our estimate of overall bias in this sector is 3.0 percent for appliances, 4.0 percent for radio-TV, including VCR's and camcorders, and 15 percent per year for personal

computers.<sup>49</sup> Applying respective current nominal weights of 0.8 percent for appliances, 1.0 percent for consumer electronics, and 0.4 percent for personal computers, this category contributes an annual rate of quality change and new product bias of 0.10 percent per year to the total CPI. The figure entered into Table 2 for this category is a weighted average of the bias estimate, but the bias figure for the total CPI is based on weights corresponding to current nominal expenditures, not the CPI weights displayed in Table 2. Also, prior to 1994 the bias figure is based only on appliances and radio-TV, since personal computers did not emerge as a significant product until that date.

Regarding housefurnishings other than appliances and video-audio products, there is no available research to provide guidance. The available range of furniture, draperies, etc., allows consumers to substitute among products, fabrics, and outlets along dimensions that are not captured by the CPI. There have been many new products in this area, including furniture and fabrics that are much less susceptible to damage by stains and childrens' accidents than was previously possible. This category also includes soap and cleaning products, where substantial progress has been made. We view a bias rate of 0.33 percent per year, or 10 percent over the past 30 years, as conservative.

3. **Apparel.** It is often assumed that there has been no quality change in apparel. But new apparel products are constantly introduced that improve consumer welfare, including denim jeans and shorts, advanced varieties of running shoes, iron-free synthetic fabrics, and lightweight but water-resistant raingear. Despite this, apparel is the other major area where the CPI is thought to have incorporated a downward bias. One source of downward bias occurred when the CPI price quotations followed the decline in price of an old model placed on sale, and then (using the deletion technique) made the transition to a new model without accurately recording the corresponding increase in price. Reforms in the CPI in the mid-to-late 1980s eliminated this source of downward bias and shifted to the hedonic price technique for some quality adjustments within the apparel component.<sup>50</sup>

The CPI apparel index is relatively easy to assess by accumulating outside evidence from such sources as mail-order catalogues. While style changes in fashion goods are frequent, quality changes in utilitarian apparel products purchased by average urban consumers are sufficiently infrequent to allow careful price comparisons across identical models from mail-order catalogues. By limiting itself to a month-to-month measurement framework, without cross-checks based on yearly or decadal comparisons, the CPI is vulnerable to persistent drift that emerges from measurement flaws such as the treatment of products on sale, as discussed above.

In a new project Gordon (1996) has compiled an apparel price index from the Sears catalogue based on thousands of year-by-year comparisons of identical apparel items over the intervals 1965-93.<sup>21</sup> The ratio of the CPI relative to the Sears apparel index rose at an annual rate of +1.92 percent per year during 1985-93.<sup>22</sup> The rapid rate of increase of the CPI apparel index after 1985 relative to Sears is surprising, because Sears in those years was losing market share to Wal-Mart and other discounters. Thus there is reason to think that the Sears catalog index might overstate the increase in true apparel prices faced by the average American consumer. Nevertheless, we shall take the conservative approach of cutting the implied bias rate from +1.92 percent suggested by the Sears index to a smaller 1.0 percent bias rate.

**4. Transportation.** The transportation component of the CPI consists of a wide variety of heterogeneous goods, including new vehicles, used vehicles, motor fuel, vehicle repairs, auto insurance and registration, and public transportation, mainly airline fares. The most important questions to be addressed in the transportation sector are the valuation of mandated safety and anti-pollution devices, and the treatment of used cars relative to new cars. As documented by Gordon (1990, p. 364) for the period 1947-83, the actual price of new cars increased much faster than the CPI for new cars, and after 1967 almost none of this relative increase could be explained by increases in the dimensions included in the traditional hedonic regression equations for new cars. The key ratios of 1983 to 1967 prices were that actual prices had increased by a ratio of 289.9, the CPI for autos had increased by a ratio of 202.6, and that the difference had been more than explained by the contribution of CPI adjustments for safety and

environmental quality and Gordon's adjustment's for fuel economy.<sup>33</sup> The resulting upward bias in the CPI relative to Gordon's final auto index is 0.44 per year from 1967 to 1983.

However, Gordon accepted the CPI's treatment of anti-pollution devices as a quality improvement rather than a price increase. We are persuaded that mandated anti-pollution devices are analogous to an indirect tax. Gasoline taxes may be used to provide a benefit in the form of better highways, but a tax increase is treated correctly by the CPI as an increase in the cost of living. Anti-pollution devices provide a benefit in the form of cleaner air, but by analogy to taxes should be treated as an increase in the price of the car. Using the detailed information given by Gordon, we can calculate an alternative to his index that converts the CPI adjustment for anti-pollution devices from a quality change to a price change, and this results in the finding that the CPI for new cars was downward biased during 1947-83 by 0.94 percent per year. We do not make a similar adjustment for the value of quality change taking the form of safety devices such as seat belts and crash-resistant bumpers, since our feeling is that consumers see the connection between their own safety and the devices more directly than they do between anti-pollution devices and air quality. Subsequently we will adjust the 0.94 percent downward bias for an offsetting increase in automobile durability.

What has happened since 1983, the terminal year of Gordon's study? Berry, Kortum, and Pakes (1996) show that emission control standards for automobiles became markedly tighter in two stages, 1975 and 1979-81, but did not change thereafter through the conclusion of their study in 1990. They also develop a fuel efficiency index that adjusts for changes in the horsepower and weight of cars; this rises by 67 percent between 1972 and 1983 and then drifts down by 13 percent by 1990. An alternative study of fuel economy (Sykuta, 1996) extends the data to 1994 and concludes that "new car fuel efficiency reached a plateau by 1986 and has not since changed by more than 1 percent in any direction" (p. 12). He attributes the absence of a decline in fuel economy in response to lower fuel prices to the high cost of "switching back to older, less efficient designs and technologies." Thus it appears that

neither changes in anti-pollution equipment nor in fuel economy have been important potential factors in creating a bias in the new car CPI since the mid-1980s.

However, neither Gordon's study nor the CPI have taken into account the increased service lifetime of the typical new car. The average age of automobiles in use increased at an annual rate of 2.1 percent per year during 1970-83, part of the period covered by the Gordon study, and at 1.3 percent per year during 1983-93.<sup>34</sup> The CPI should be pricing automobiles on a rental equivalent basis, parallel to their treatment of housing. If the useful life of a car is improved by technological change that raises quality, reduces maintenance requirements, and minimizes rusting and corrosion, then consumers benefit. The cumulative increase in the useful life of a car from 1970 to 1993 amounted to 48 percent. Consider an automobile costing \$10,000 with a life of 10.0 years; this is equivalent to an annual cost of \$1,000 per year. Now consider the same car with a life extended by 48 percent to 14.8 years; its annual cost has been correspondingly reduced to \$676.

The reduction in automobile depreciation is only part of the user cost of owning a car. Many of the additional elements of cost, e.g., insurance and repairs, are priced separately by the CPI in the major category "Other Private Transportation," where we impute no bias at all. The remaining component of user cost, in addition to depreciation, is real interest expense. Balancing alternative methods of paying for cars, including cash, installment finance, and leasing, we think that 10 percent is a reasonable estimate of the real interest cost. The extension of automobile longevity has reduced user cost from roughly 20 percent per year (consisting of a 10 percent real interest cost plus 10 percent depreciation rate) to 16.7 percent (the same real interest cost plus a 6.7 percent depreciation rate), for an overall reduction in user cost of 16.7 percent, or a geometric rate of -0.79 percent per year over the period 1970-93.

Since the rate of increase of average age slowed after 1983, we distribute this effect accordingly, as contributing to an upward bias in the CPI at a rate of 0.95 percent per year during 1970-83 and 0.59 percent per year after 1983. For the 1970-83 period, we add the upward bias in the true price of cars due to increasing durability of 0.95

percent per year to our previous estimate of a downward bias of 0.94 percent per year, exactly cancelling out and yielding a net bias for zero. For the subsequent post-1983 period, we found no other reason to suppose the CPI is biased in either direction, so the durability adjustment is converted to a net upward bias of 0.59 percent per year.

In assessing these calculations, we recognize that some fraction of the durability effect may not represent a pure quality change, but rather may reflect other factors, such as an increased relative price of cars that induces users to hold onto cars longer, or an improved quality of highways. This might suggest that our durability adjustment is too large. However, an offsetting error may be more important and may represent an important source of quality improvement that is not taken into account in the CPI or in our adjustments -- namely, the marked decrease in the incident of defects on both imported and domestic new cars, as measured by the J. D. Power survey and other evidence. This additional source of quality improvement, which we have not taken into account, suggests that our estimate of the CPI bias for new cars is probably conservative.

The CPI index for used cars has long been known to be upward biased, simply because no quality adjustments were applied to this category at all. The upward bias over the period 1967-87 is very large, amounting to 2.44 percent per year, if we take the difference in the growth rate between the new car and used car index to represent a measure of the bias.<sup>22</sup> Adding this to the new car upward bias of zero, we arrive at a total bias for used cars for 1967-87 of 2.44 percent per year. For the period since 1987 we apply the CPI durability bias to used cars as well.

Regarding other components of the transportation category, we regard motor fuel as homogeneous and easy to measure. However, numerous improvements in fuels and related products have contributed both to cleaner air and improvements in autos, such as increased durability. While the cleaner air is valuable in its own right (see Section VII), we treat the mandate for cleaner gasolines like an indirect tax, not quality change per se. The improvements to autos are counted there, but must be understood to result in part from improvements in fuels, lubricants, etc. The CPI treats full-service and

self-service fuel as two different commodities, as in the case of full-service department stores and no-frills discount stores. The shift in market share from full-service to self-service motor fuel provides evidence of a type of outlet substitution bias, which we do not count here as quality change bias. For the past decade, when the transition to self-service was largely complete, we attribute a small upward bias of 0.25 percent per year to the CPI for ignoring the increased convenience and time-saving contribution of automatic credit-card readers built into gasoline pumps.<sup>26</sup>

"Other private transportation" expenditures, an important CPI category, consist primarily of automobile maintenance and repair and automobile insurance. We are not aware of any evidence of price index bias involving these areas and thus assign a zero bias. However, in the public transportation category, which is dominated by airline travel, it is well documented that prior to 1982 the CPI incorporated a substantial upward bias due to a failure to take discount fares into account. We take our estimate of this bias from Baily-Gordon (1988, p. 416) and multiply it by the weight (about 2/3) of airline fares in the public transportation category.

**5. Medical Care.** The medical care category may be the location of substantial quality change bias at a rate as rapid or more rapid than in appliances and radio-TV products, but its weight in the CPI is much greater. That weight is controversial in itself, since the 7.4 percent weight is based only on out-of-pocket expenses by consumers and does not include expenses paid for medicare, medicaid, or by employer-financed payroll deductions. Since one of the main uses of the CPI is to convert indexes of real income, compensation, and wages into "real terms," the current weight is wrong. Changes in medical care prices or technology should influence the CPI with the same weights that are relevant for total income and compensation, that is, they should include the total medical care bill, which in National Income and Product Accounts version of consumption amounts to 16 percent, not the much smaller 7.3 percent weight in the CPI.

The CPI weight of 7.4 percent is applied to three primary categories -- drugs, professional medical services (i.e., doctors and nurses), and hospitals. There has been

substantial recent research on prescription drugs. For instance, Berndt, Cockburn, and Griliches (1996) have studied prices of anti-depressant drugs. They find a substantial difference between an index based on BLS methods, which rises during 1993-96 at 4.6 percent per annum, and an alternative index based on their preferred methodology which rises at 1.1 percent per year. This alternative index uses an alternative (Divisia) weighting scheme, introduces new items much more promptly, and treats generic drugs as fully equivalent to proprietary brand-name drugs. In July, 1995, the CPI shifted its treatment of generics, so that any decline in price when generic equivalent drugs become available is taken fully into account as a price change rather than being treated as a different good and hence "linked out" of the index. Based on the Berndt *et. al.* research, and related research by Griliches and Cockburn (1994), we conclude that prior to 1995 there was a 3.0 percent annual bias in the CPI for prescription pharmaceuticals. The new CPI treatment of generic drugs after July 1995 reduces this annual rate of bias from 3.0 to 2.0 percent, and the remaining bias can be attributed (based on the previously cited research) to the late introduction of particular drugs into the CPI and the failure to attribute a positive value to newly introduced drugs that gain market share and thus appear to add value for consumers.

The major weight of the CPI medical care component is applied to medical care services, e.g., physician fees, and hospital costs. No attempt is made in the CPI to value health care "outcomes" as contrasted to "inputs". A hybrid approach is taken by the Producer Price Index (PPI), which also prices inputs but with different weights and increases by roughly 2.0 percent per year more slowly than the CPI in both the doctor and hospital category in the recent past (1995-96). Thus government indexes already provide important evidence that the CPI is upward biased by at least two percent per year, but that ignores many changes in medical care practice and technology that suggest a higher rate of bias. Cutler *et. al.* (1996, Table 1) contrast input price indexes for the cost of heart-attack treatment with alternative "outcome " indexes that take account of the cost of achieving a given increase in life expectancy. The authors contrast an "input" index of the type currently compiled by the CPI, which rises by 3.3 percent per year over 1983-94, with a final outcomes index that takes into account a conservative valuation for the extension of life expectancy attributable to new heart

attack treatments and *declines* by 1.1 percent a year, for a net bias of 4.4 percent per year. Shapiro and Wilcox study cataract surgery for the period 1969 and 1993. Their "prototypical" index that duplicates CPI methods increases by a factor of 9 while their preferred alternative index increases by a factor of 3, implying an annual rate of upward bias of 4.57 percent. The closeness in the Cutler *et. al.* and Shapiro-Wilcox studies of quite different medical procedures is striking. In view of the fact that the CPI has been rising relative to the PPI for medical care (taking the input rather than the outcomes approach) by 2 percent per year in 1995-96, the rate of upward bias that we have assigned to the medical care category, 3.0 percent per year, is probably conservative. There is probably additional quality change taking into account the wide variety of new diagnostic and test equipment, the reduction of pain of routine procedures, the shortening of hospital stays, and also the quality of hospital rooms.

This new research by Cutler and others opens up the potential for a major improvement in our understanding of the economics of medical care. This category should receive a substantial component of the CPI's future research investment, and we strongly endorse a move in the CPI away from the pricing of health care inputs to an attempt to price medical care outcomes.<sup>24</sup>

**6. Entertainment.** The entertainment category is divided roughly equally between commodities and services. Commodities consist of newspapers, magazines, sporting equipment, and toys. We assume that there is no bias in newspapers and magazines (although electronic news services provide convenience and timeliness to some) but that sporting equipment and toys are subject to a somewhat smaller bias than Gordon found for appliances, namely 2.0 percent per year as contrasted to 3.0 percent for appliances. This may represent an average of the rate of bias in electronic toys, e.g., Nintendo games, which may be close to the 15 percent rate we have applied to personal computers, and a bias rate of zero for other toys, including stuffed animals and non-electronic board games. Entertainment services consist of club memberships, admission fees, and lesson fees. There has undoubtedly been an improvement in the comfort of attending sports events, with domed stadiums and more comfortable seats, but we have not assigned any bias to the CPI measure of entertainment services prices.

**7. Other Goods and Services.** While purchases of such products as cigarettes, toiletries, and cosmetics may have been subject to outlet substitution bias, there is no reason to think that quality change has been missed or that new product bias has been important. However, this category includes a small weight for small personal care appliances, e.g., hair dryers, and it is reasonable to suppose that these items are subject to the same magnitude of bias as large appliances, i.e., 3.0 percent per year. Much of the rest of the weight in the "other" category consists of school books and fees, primarily college tuition. However, about one-tenth of the weight in this category consists of personal financial services, which have been subject to rapid technological change, particularly with the widespread diffusion of ATM machines and all-in-one cash-management accounts. Taking a conservative 2.0 percent estimate of personal financial services and applying that to one-tenth of this category yields our 0.2 percent annual rate of bias.

### *Conclusion on Quality Change Bias*

Our final estimate of quality change bias, taking the weights and bias rates in Table 2, aggregates to approximately 0.6 percent per year. This is slightly higher than the rate of 0.5 percent per year for the combined categories of quality change and new product bias estimated in our interim report or 0.35 percent in Shapiro and Wilcox (1996c). Our higher estimate results from a much more extensive examination of the existing literature and consideration of factors that affect each of the 27 separate categories included in Table 2.

TABLE 2: Relative Importance Of Components In The CPI-U, U.S. City Average, December 1995, And Estimates Of Quality Change And New Product Bias For Selected Time Intervals				
Major and Selected Minor Components	Relative Importance in Percent	Estimated Quality Change Bias at annual rate for Selected Time Intervals		
I. Food and beverages	17.332			
Food at home other than produce	8.543	0.30 (1967-96)		
Fresh fruits and vegetables	1.337	0.60 (1967-96)		
Food away from home	5.886	0.30 (1967-96)		
Alcoholic beverages	1.566	0.15 (1967-96)		

2. Housing	41.346			
Shelter	28.289	0.25 (1976-96)		
Fuels	3.792	0.00		
Other Utilities, incl. telephone	3.222	1.00		
Appliances incl. electronic	0.806	3.6 <sup>a</sup> (1973-94)	5.6 <sup>a</sup> (1994-96)	
Other housefurnishings	2.639	0.30		
Housekeeping supplies	1.116	0.00		
Housekeeping services	1.482	0.00		
3. Apparel and upkeep	5.516	-0.95 (1965-85)	1.00 (1985-96)	
4. Transportation	16.953			
New vehicles	5.027	0.00 (1970-83)	0.59 (1983-96)	
Used Cars	1.342	2.44 (1967-87)	1.59 (1987-96)	
Motor Fuel	2.908	0.0 <sup>b</sup> (1974-84)	0.25 <sup>b</sup> (1984-96)	
Other Private Transportation	6.153	0.00		
Public Transportation	1.523	2.66 (1972-77)	4.60 (1977-82)	0.00 (1982-96)
5. Medical Care	7.362			
Prescription drugs	0.891	3.00 (1970-95)	2.00 (1995-96)	
Nonprescription drugs and medical supplies	0.391	1.00		
Professional medical services	3.465	3.00		
Hospital and related services	2.257	3.00		
Health insurance	0.358	0.00		
6. Entertainment	4.367			
Commodities	1.975	2.00		
Services	2.392	0.00		
7. Other Goods and Services	7.123			
Tobacco, smoking products	1.610	0.00		
Personal care	1.170	0.90		
Personal and educational expenses	4.342	0.20		
<b>Total</b>	<b>100.000</b>	<b>0.612 (for 1996)</b>		

Notes: <sup>a</sup> Applied to the weight of this category in 1995 annual nominal personal consumption expenditures, 2.13 percent, as contrasted to the December, 1995 CPI relative importance of 0.8 percent.

<sup>b</sup> See text.

Has quality change bias increased or decreased? Table 2 provides some insights to that issue. To take the specific example of 1980, we can aggregate the rates of bias given in Table 2 and come up with an upward CPI bias due to quality change and new products of 0.488 percent per year, compared to 0.613 for 1996. The main differences come from the reversal of the previous downward bias for apparel and the increased upward bias in the appliance/radio/TV component due to the growing role of personal computers. Partially offsetting these sources of increased upward bias are reductions in the extent of upward bias in used cars, airline fares, and prescription drugs.

## **VI. Estimates Of Biases By Type And In Total**

The CPI is not a cost of living index, but rather a fixed weight index, implemented by means of a modified Laspeyres formula. This creates an immediate conflict between the objectives of the CPI and many of the purposes for which it is intended. For example, the CPI is used to index private contracts, tax brackets and government transfer programs, such as Social Security, in order to compensate beneficiaries for changes in the cost of living. A fixed weight index exaggerates the effect of price changes on the cost of living, because it fails to allow for substitutions that enable consumers to avoid the full impact.

The assessment of biases in the CPI requires a cost of living index as a point of reference. An approximation to a cost of living index can be generated by combining the results of research on different types of biases in the CPI. The purpose of this section is to summarize this research and assess the importance of these biases in the CPI. Our estimates depend primarily on studies produced prior to the convening of our Commission, many of them by the BLS.

Our Interim Report of September 15, 1995, presented initial estimates of biases in the CPI. We estimated that the overall bias had been 1.5 percentage point per annum in recent years, but changes in the CPI methodology then in prospect from BLS would eliminate as much as 0.5 percentage point per annum of this bias, reducing the bias going forward to 1.0 percentage point per annum. We have now revised our estimates

to reflect changes in the CPI announced by BLS on March 29, 1996, and new estimates of the impact of biases due to the introduction of new products and changes in the quality of existing products. The BLS has eliminated some of this bias totaling 0.24 percentage point per year, raising our estimate of the bias going forward by one-quarter of one percentage point. In addition, we have revised our estimates of new products/quality change bias upward by 0.10 percentage point per year.<sup>58</sup>

In assessing biases in the CPI it is essential to separate two types of substitution bias. First, BLS uses a fixed weight index based on the modified Laspeyres formula to combine price indexes for 207 items for 44 areas into a national CPI. The weights are derived from the Consumer Expenditure Survey (CES) and reflect surveys of individual households. We refer to the substitution bias at this level as Upper Level Substitution Bias. This bias is measured as the difference between the modified Laspeyres formula used by BLS and a Tornqvist index, which is (approximately) free of substitution bias.<sup>59</sup> Most estimates cluster around 0.2 to 0.25, including numerous estimates from BLS. The latest estimates available to the Commission reflecting just produced unpublished corrections of previous research by BLS show an average bias over the period 1988-1995 of about 0.15 percentage point per year. While we have not had time to analyze these new results, to be conservative, we adopt this figure.

The second type of substitution bias results from combining price observations for approximately 71,000 goods and services and information on prices for housing components of the CPI into indexes for the 207 items and 44 areas. We refer to the substitution bias at this level as Lower Level Substitution Bias. The prices to be collected are selected by probability sampling. The probabilities are derived from the CES and the Point of Purchase Survey (POPS) of retail establishments and are intended to reflect the share of items and areas in consumer expenditures in the base period.

Beginning in 1978 BLS introduced a sample rotation procedure in the estimation of price indexes within each item and area category. At the same time prices were combined in a way that reflected the modified Laspeyres index formula.<sup>60</sup> These two

changes had the effect of introducing considerable Lower Level Substitution Bias, a fact discovered by Reinsdorf (1993) after a lapse of fifteen years.<sup>o1</sup> This issue has been intensively studied by BLS and steps have been taken to deal with it, beginning in January 1995.<sup>o2</sup>

For data from the CPI covering the thirty month period from June 1992 to December 1994, Moulton and Smedley (1995) have estimated that the difference between the modified Laspeyres formula used at the lower level of aggregation by BLS and a weighted geometric mean formula for non shelter components of the CPI to be 0.49 percentage point per annum. This difference is an estimate of the bias of the Laspeyres formula, since Shapiro and Wilcox (1996) have shown that the geometric mean provides an unbiased estimate of the underlying cost of living index. Armknecht, Moulton, and Stewart (1995) have estimated the bias for owners' equivalent rent to be 0.50 percentage point per annum.<sup>o3</sup> BLS is currently testing the feasibility of producing indexes based on this methodology for price data beginning in 1990. Substitution of the weighted geometric mean formula for the modified Laspeyres formula at the lower level of aggregation is an important step in the direction of a cost of living index.

In January 1995 the BLS introduced an improved method for the imputation of price changes for food at home, owners' equivalent rent, and prescription drugs. These changes are described by Armknecht, Moulton, and Stewart (1995). For food at home items the changes included the introduction of a procedure called "seasoning." The seasoning period is the time to obtain the data needed to weight each new sample observation before introducing it into the index. This period was lengthened to three months for food at home items, breaking the link between the weights for these items and prices eventually used in the CPI.

On March 29, 1996, the BLS announced that the seasoning procedure would be extended to all non shelter items, effective with the CPI for June 1996.<sup>o4</sup> The announcement pointed out that residential rent and owners' equivalent rent were no longer subject to the bias associated with sample rotation procedures, as a consequence

of the changes introduced in January 1995. In addition, the BLS stated that (with rare exceptions) the weight for a substitution item would be kept constant throughout the life of the item. BLS estimated the reduction in bias due to the January 1995 changes to be 0.14 percentage point per annum, while the 1996 changes for the non-shelter items reduced the bias by a further 0.10 percentage point per year. This reduces the Moulton-Smedley estimate of the remaining Lower Level Substitution Bias to 0.25 percentage point per annum. On July 16, 1996, BLS introduced changes in the classification and definition of the hospital and related services component of the CPI.<sup>65</sup> These were intended to improve the measurement of this important component of the CPI, but were not accompanied by an estimate of bias reduction. Thus, our point estimate of Lower Level Substitution Bias is 0.25 percentage point per annum.

The goal of BLS is to measure goods and services of constant quality; however, the disappearance of products from the marketplace necessitates the substitution of other products. Armknecht and Weyback (1989) summarized the methods used by BLS to adjust the CPI for quality change.<sup>66</sup> Elimination of new goods bias and quality change bias are essential steps in measuring the cost of living. Important empirical research on new goods bias has been done for breakfast cereals by Hausman (1996), prescription drugs by Griliches and Cockburn (1994), new cars by Pakes, Berry, and Levinsohn (1993), and many others as described above in Section V.<sup>67</sup> We have estimated the total bias due to new products and quality change of existing products to be 0.6 percentage point per year.

Reinsdorf (1993) has provided the principal empirical evidence on New Outlet Substitution Bias. This is based on comparisons between prices for certain food and fuel items for outlets rotating into the sample covered by the CPI and outlets rotating out. He estimated the bias to be 0.25 percentage point per annum. Lebow, Roberts, and Stockton (1994) have extrapolated this estimate to the CPI as a whole by identifying components of the index that would be affected by outlet substitution bias. These amount to 40% of the CPI, so that Outlet Substitution Bias is 0.1 percentage point per annum.

These separate biases are approximately additive and likely to be independent of modest swings in the true inflation rate. Thus, a bias of 1 percentage point implies that when changes in the CPI show inflation rising from 3% to 5%, it is likely actually to be rising from 2% to 4%. Note the bias primarily affects the level, not the change, in the inflation rate. At very high rates of inflation, the bias may increase (one might assume greater outlet and commodity substitution), but we currently have no evidence regarding this issue.

Table 3 summarizes our evaluation of biases in the CPI. This includes point estimates based on the best available evidence as well as a plausible range for the overall bias. The average of our estimates of the overall bias in the CPI is 1.1 percentage point per annum with a range of 0.8 to 1.6 percentage point.

<b>TABLE 3: Estimates Of Biases In The CPI-Based Measure Of The Cost Of Living (Percentage Points Per Annum)</b>	
<i>Sources of Bias</i>	<i>Estimate</i>
Upper Level Substitution	0.15
Lower Level Substitution	0.25
New Products/Quality Change	0.60
New Outlets	0.10
<b>Total</b>	<b>1.10</b>
Plausible range	(0.80-1.60)

The BLS is preparing for a benchmark revision in January 1998, when the CPI will incorporate new expenditure weights from the 1993-1995 Consumer Expenditure Surveys. However, BLS will retain the modified Laspeyres formula, so that our estimates of bias will carry over to the revised CPI. In addition, BLS has continued to introduce important modifications in the CPI in order to improve measurements and remedy deficiencies that have come to light. However, these revisions, like the forthcoming benchmark revision, employ the modified Laspeyres framework, so that important differences between the CPI and a cost of living index will remain.

The Upper Level Substitution Bias in the CPI will persist after the forthcoming benchmark revision of the CPI, since BLS plans to retain the modified Laspeyres formula. Second, BLS has reduced so-called formula bias, the part of Lower Level

Substitution Bias resulting in substantial measure from the introduction of sample rotation procedures and the modified Laspeyres index at the Lower Level of aggregation in 1978. However, the objective of the changes in January 1995 and those announced in March 1996 was to improve the implementation of the modified Laspeyres formula, not to eliminate the Lower Level Substitution Bias quantified by Moulton and Smedley (1995). Finally, New Item, New Outlet, and Quality Change Biases are unaffected by the changes already announced by BLS or the benchmark revision.<sup>68</sup>

To summarize we have revised the estimates presented in our Interim Report to reflect BLS revisions of the CPI and the accumulation of new findings on the magnitude of biases. Our main conclusion is that the limitations imposed by the modified Laspeyres formula make the CPI unsuitable for cost of living measurement. By combining a Tornqvist formula at the Upper Level of aggregation and a weighted geometric formula at the Lower Level, BLS could eliminate both types of Substitution Bias. However, these changes alone would fail to encompass adjustments for New Item, New Outlet, and Quality Change Biases. Adjustments for these biases are essential for measurement of the cost of living.

Figure 2 illustrates the compounding effect of a 1.1 percentage point per annum bias on CBO projections of the CPI U through 2006. While 1.1 percentage point may seem to be a small amount in any given year, cumulatively year after year it adds up to a sizable difference, 14% over a dozen years.

## **VII. Other Issues**

### ***A. A Separate Price Index for the Elderly?***

In principle, if not practice, a separate cost of living index could be developed for each and every household based upon their actual consumption basket and prices paid. As noted above, the aggregate indexes use data reflecting representative consumers. Some have suggested that different groups in the population are likely to have faster or slower growth in their cost of living than recorded by changes in the CPI. We find no

compelling evidence of this to date, and in fact two studies suggest that disaggregating by population group, for example by region or by age, would have little effect on measured changes in the cost of living.<sup>99</sup> Further, work on this subject remains to be done. In particular, the prices actually paid, not just expenditure shares, may differ.<sup>100</sup>

Beyond the different consumption baskets, it is important to understand our analyses of the sources of bias are applied to representative or average consumers. Some consumers will substitute more than others, and the substitution bias may be larger for some, smaller for others. Likewise, some are more likely to take advantage of discount outlets; others less so. Perhaps most importantly, the benefits of quality change and the introduction of new products may diffuse unevenly throughout the population. Some will quickly gain the benefits of cellular telephones, for example, while others may wait many years or decades or never use them. This is yet another reason why we have been very cautious in our point estimates for these particular sources of bias.

### ***B. Broader Considerations on the Quality of Life***

Not all change is positive and not all change is positive for everyone. In making the case for the importance of quality change, we have also to look at the other side of the ledger. There are at least three types of change to consider here: 1. New goods may drive out older goods which are still valued by a subgroup of the population, or what is equivalent, the loss of economies of scale may drive up their price significantly. To the extent that it is measured, it does not represent a new problem of the price index construction. 2. An existing good or service may deteriorate in its quality. That is a less frequent phenomenon, in spite of the mantra that "they don't make them the way they used to." The most significant recent example is probably the HMO-induced tightening-up of the rules of access to medical care which is likely to be perceived as a significant deterioration in the expected services that had been contracted for by the purchase of medical insurance. It is unlikely, however, to have outweighed the medical advances of recent years. Few would trade today's restricted access to medical care, for a more free access to the technologies of yesteryear, foregoing the improvements in bypass operations, ulcer treatments, or cataract surgeries. 3. The largest effect may come from change in our physical, social, and economic environment which impose

on us higher expenditures necessary to keep up with our previously achieved utility levels.

It is not clear, however, whether events such as a colder winter, the appearance of AIDS, or a rise in the crime rate should be included in the definition of a *price* index. A change in expenditures due to an unanticipated change in weather should raise the price index only to the extent that energy prices go up, not quantities consumed. The latter, if the event persists, will ultimately affect the commodity weights in the index, but that is a different matter. The rise in AIDS would drive up the price index of health, if we define it as the expenditure necessary to achieve an equivalent base-period health level. But while this component represents a real rise in the "cost-of-living," it may not be an appropriate component in an indexing formula, since there are no "gains" among the young which could be called to "compensate" the retirees for such price increases.

The appearance of AIDS did raise the cost of living. Counting subsequent improvements in its treatment as a positive quality change will be inappropriate, if the original deterioration in the environment was not taken into account in the measurement framework in the first place. Similarly, counting quality improvements in locks and other security devices, may overestimate the "gains" from such defensive consumer investments.

While it is impossible to provide a full and accurate accounting for such changes, it is extremely unlikely that the rise of such "bads" out-balances the "good" that is contained in the developments alluded to above. In the major areas of concern and poor measurement, environment, health, and crime, there have been gains in the first two and we have come close, in recent years, to holding steady on the last one. Thus, while we do recommend extending measurement efforts beyond the current concept of what constitutes the consumption "basket," we see no strong reason to temper our conclusions about the measurement issues in the areas we did examine: the more traditional components of consumption as defined in the current content of the CPI.

The following presents some brief background for our view.

The industrial revolution caused widespread air and water pollution, and this was indeed a negative factor up until the 1950s and 1960s. However, since then a shift from coal to natural gas as the dominant energy source for home heating, cleaner fuels and cars and environmental regulation have caused a major decline in the presence of many types of contaminants in the air and water. The shift in heating fuels also brought about a major increase in cleanliness and convenience around the house, as the coal scuttle was replaced by the silent and automatic transmission of natural gas. The CPI implicitly values the improvement in air quality made possibly by mandated anti-pollution devices in automobiles, since it treats the cost of mandated anti-pollution devices as an improvement in quality rather than an increase in price. We have recommended that the CPI treat changes in price due to additional anti-pollution devices as a price rather than quality increase. But that concept, incorporated in our quality change bias estimates in this report, leaves the improvement in air quality unmeasured. This is a source of understatement in our final estimate of CPI bias. Further, the CPI is inconsistent, since a portion of the higher cost of electricity, steel, paper, and other products is also due to environmental regulation, and the benefits of higher air and water quality made possible by regulation of products other than automobiles is not taken into account.<sup>11</sup>

Crime is another type of externality. The rate of crime victimization increased in the past, e.g., from 0.096 incidents per capita in 1970 to 0.098 incidents in 1981. Since then, however, the victimization rate has fallen to 0.074 incidents in 1992 (the latest year available).<sup>12</sup> The share of violent crimes increased from 0.025 in 1970 to 0.028 in 1981 and decreased slightly to 0.026 in 1992. Since 1992 there has been widespread attention to a sharp decline in homicide rates in many major American cities.

Looking further for negative factors, perhaps the most important are such social issues as divorce, illegitimacy, and the reduced role of the nuclear family. The divorce rate increased by 50 percent between 1970 and 1980 but since then has been stable.<sup>13</sup> The suicide rate was stable between 1980 and 1992 but the age-adjusted death rate declined by 14 percent, while the infant mortality rate fell by 58 percent between 1970 and

1992. Perhaps most importantly, life expectancy at birth increased from 70.8 in 1970 to 75.5 in 1993. The major negative has been that births to unmarried women have increased, from 18.4 percent to 30.1 percent. There may be other intangible negatives, such as perceived increased job insecurity, possible increased inequality, and the decreased job opportunities for workers with only a high school education.

On the positive side, there is no question that goods have improved in ways that our discussion of quality change cannot fully incorporate. Gordon (1990, p. 38) notes a number of dimensions of quality which his measures did not capture, some of which are the faster speed and reduced vibration of jet planes, improved reliability of appliances and automobiles, improved sound quality of audio equipment in homes and automobiles, improved safety devices on home power tools and power lawn mowers, reduction in the noise, weight, and installation cost of room air conditioners, and "immeasurably better picture quality of color TV sets."

Many aspects of the change in the American standard of living from the rural horse-drawn economy of 1870 to today's modern economy occurred many decades ago and are not current sources of CPI bias. The elimination of animal waste was a major contribution of the motor car, in addition to its speed and flexibility, but this achievement was largely accomplished before World War II. The achievement of electrical appliances in reducing household drudgery was largely accomplished by the 1960s. But some improvements have been continuous. There has been a steady transition in the quality and variety of home entertainment options, from the player piano, to the radio, to black-white and then color TV, to the VCR, and now cable TV, satellite TV, and the World Wide Web and the other features available with personal computers. Further, the rapid spread of central air conditioning has made possible a substantial movement of households to the southern and southwestern states. Millions have chosen to take this option, voting with their feet to enjoy milder winters with artificially cooled summers.

Overall, we find that the presumed negatives (pollution, crime, suicide, divorce), the worsening of which may have increasingly detracted from the quality of life at one

time, have reached a plateau and in the case of pollution and crime seem actually to have reversed direction, thus recently improving the quality of life. The remaining negatives are important but seem to us to have been more than offset by increased quality and variety of goods, services, and choice of outlets along dimensions that are partly but not entirely captured by our measures of bias, but most importantly by the major increase in longevity which perhaps swamps everything else. Accordingly, our estimate of the current bias in the CPI is, if anything, probably understated.

## **VIII. The Commission's Recommendations**

### **Introduction**

Our recommendations are directed at three different audiences: 1. Our elected officials (the President and the U.S. Congress), who are the providers of funds, the supervisors, and also major consumers of the information contained in the CPI; 2. The BLS, which is the producer of the CPI; and 3. The community of professional economists and statisticians, who ultimately provide both the manpower and the knowledge base required for a successful operation of this major measurement and observational tool for our economy. The role of the BLS in this can be viewed as running one of our major National (Economic) Observatories, producing both timely information about the state of our economy and providing the inputs for advancing our knowledge of and understanding about how it functions and the interrelationships among its various components.

### **Recommendations**

**1. The BLS should establish a cost of living index as its objective in measuring consumer prices.**

All of our specific recommendations are aimed toward this goal.

## **2. We recommend the development and publication of two indexes:**

One which is published monthly on a timely basis and is designed to maintain the spirit of the cost of living index yet accommodate the inconsistent timing schedules of the required information; and a second index which is published and updated annually and revised historically to introduce improvements arising from new information and new research results. The purpose of having two indexes is to accommodate the complex issues that must be addressed and the time delay in obtaining all of the necessary data.

The recommendations to the BLS are divided into three parts: 1. Short run: those we think can be implemented immediately, with little additional resources or new data collection initiatives. These center on changing the current CPI computation, primarily to make it more current, and second, on computing an annually updated and subsequently revised COL index; 2. Intermediate run: Reforms that are feasible within the current state-of-the-arts, but would require new data collection, reorganization of activities, and changes in the detail of the various sub-indexes produced by the CPI; and 3. Longer-run recommendations, emphasizing topics and areas that need additional research and attention.

### ***Short Run***

**3. The timely, monthly index should continue to be called the CPI and should move toward a COLI concept by adopting a "superlative" index formula to account for changing market baskets, abandoning the pretense of sustaining the Laspeyres formula.**

To accommodate the delay in obtaining information on quantities needed to combine the price changes of items in the lowest groupings, BLS should move away from the assumption that consumers do not respond at all to price changes in close substitutes. We thus recommend BLS move to a "trailing Tornquist" Index (weighted geometric mean of price relatives), at the stratum and ELI level, and also, concurrently, to

geometric means of price relatives at the elementary aggregation level. Both of these moves would alleviate the problem of the growing irrelevancy of "baskets" based on decade-old consumption patterns, reduce significantly the substitution and formula bias, and facilitate the speedier introduction of new goods and services into the index.

/4

a. Because of the lag in collecting up-to-date information on consumer spending patterns, the weights will have to be based on a trailing two- or three-year average of past expenditures, e.g. 1993-4 weights for the 1996 price changes. They should be changed every year.<sup>15</sup> This implies that

b. The BLS should organize itself for "permanent" rather than decadal revisions in the CPI. Both the weights and the priced commodity and services assortment need more frequent updating. Also,

c. Wherever possible, scanner data and other "outside" data should be used both to reduce the cost of data collection and (primarily) to expand the assortment of goods and services priced concurrently, to provide current item weights, and to introduce new items as quickly as they enter the market. Whether this will result in a net reduction in the cost of data collection is an open question.

d. As subsequent data become available, the weights are updated, and new goods are introduced and their history extended backward, the information incorporated in the published CPI should undergo retroactive revision, as far back as the new information warrants, in the form of a new annual COL index, using a compatible "superlative-index" formula. This "revised" COL index would be published annually, with a lag of a year or two, and would be subject to additional revisions after new information emerges and new methodology is introduced. The published versions of this index need never be "final".

#### **4. The BLS should move to geometric means at the elementary aggregates level.**

We believe it to be the closest approximation to a full implementation of the COLI

concept, which could be ultimately implemented also at this level, as scanner data become available for most of the currently sampled commodities. Changing to geometric means will not only solve the "formula bias" problem, much of which has been recently eliminated by a switch to "pre-seasoning", but will also alleviate the below-stratum-level substitution bias. It will not solve, however, the "outlet bias". To aid in updating the required weights, the BLS should be able to acquire the detailed commodity-level shipments data currently collected by the Census but not accessible to the BLS.

### *Intermediate Run*

**5. The BLS should study the behavior of the individual components of the index to ascertain which components provide most information on the future longer-term movements in the index and which items have fluctuations which are largely unrelated to the total and emphasize the former in its data collection activities.**

This could result in the down-weighting or even elimination of data collection for certain cities and a revision of the commodity structure of the index which would consider some goods as having a national market, sampling a larger number of items but with less regard to geography, focusing on geographical differences only for more "local" commodities, such as fuel costs, rent, personal services, and fresh produce.

Currently, the BLS collects a large number of price quotes on bananas, because they are inexpensive to collect and their prices are quite variable, even though these variations are not related systematically to the underlying trend-movements in the CPI. At the same time, less attention is paid to less variable but more likely to change (disappear or be redesigned) and harder to measure commodities, such as surgical treatments, consumer electronics, and communication services.

**6. The BLS should change the CPI sampling procedures to de-emphasize geography, starting first with sampling the universe of commodities to be priced**

**and then deciding, commodity by commodity, what is the most efficient way to collect a representative sample of prices from which outlets, and only later turn to geographically clustered samples for the economy of data collection.**

The current city level price indexes are useless for geographical comparisons of levels and misleading as measures of rates of change, since they are not based on any clearly defined levels. To do an adequate job of describing the geography of price *levels* in the U.S. will require the collection of prices for the *same* commodities and services in different cities. To study differential *changes* in the price levels across cities, arising from different competitive and population trends, it may prove adequate to sample the "national" commodities in specific cities only once a year or so, on a rotating basis. More generally, one could design a model consisting of an underlying "national" trend level of the CPI, which would be the primary focus of monthly estimation, and more slowly changing city differentials, which would be based on less frequently collected data.

This would allow the CPI to concentrate resources on expanding the sample and analysis in rapidly changing areas of the commodity and services spectrum, such as health services, communication services, and food away from home, where quality change and commodity turnover is endemic.

Moving to a notion of a new "basket" each year will allow for a faster introduction of new items and new outlets. Moving to a national sample for most of such items would allow expansion of the number of specific items (models, varieties, types) sampled within a particular ELI and reduce thereby the number of forced substitutions. Also, this would allow for the use of new sources of data, such as scanner data on prices, and industry-wide information on sales of specific items (for more detailed weights), leading to a quicker identification of new goods and their faster incorporation into the index. This is also the level at which more extensive quality adjustments and "comparable" substitutions could be made, recognizing the appearance of new outlets and new versions of services which provide consumers, effectively, with cheaper sources for the same or similar items consumed previously.

**7. The BLS should investigate the impact of classification, that is item group definition, on the price indexes, to improve the ability of the index to fully capture item substitution.**

In addition, a classification rule should be implemented for new products that groups them within the same low-level group (stratum) as those for which consumers are most likely to substitute for them. On-line news services which compete with newspapers, automobile purchases with leases, and drugs with the surgical procedures they replace, are examples of products for which direct comparisons are needed so that the full substitution effect can be accounted for.

**8. There are a number of additional conceptual issues that require attention. The price of durables, such as cars, should be converted to a price of annual services, along the same lines as the current treatment of the price of owner-occupied housing. Also, the treatment of "insurance", should move to an ex-ante consumer price measure rather than the currently used ex-post insurance profits based measure.**

**9. The BLS needs a more permanent mechanism for bringing outside information, expertise, and research results to it.**

This commission did not have the resources or the time to investigate all the various aspects of the CPI in adequate depth. Nor would a subsequent similar group if it were again assembled ad hoc. A more permanent body should be created, at the request of the BLS, organized by an independent public professional entity, such as the American Economic Association, the NRC-NAS or NBER, with a significant resource commitment. Such a group could pursue more fundamental research in cooperation with the BLS and provide a framework for experimentation with various alternative data collection and estimation approaches. It would also provide the BLS with a more permanent channel for access to a range of professional and business opinions on the statistical, economic and current market issues arising in the normal process of data collection, on index number construction, and on the implementation of some of the reforms suggested here.

## ***Longer Run***

### **10. The BLS should develop a research program to look beyond its current "market basket" framework for the CPI.**

In the longer run, the big issues are new commodities and new services and the changing economic, social, and environmental climate within which the consumer is operating. This program should explore measuring the value of time saved by new medical procedures and communication devices, the value of life extended and its associated quality, the losses experienced, in terms of longer distances to new shopping centers by the closing of some neighborhood stores, and the "consumption" increases forced on consumers, by rising crime, new diseases, or changes in taxation.

### **11. BLS should investigate the ramifications of the embedded assumption of price equilibrium.**

This assumption, which means that prices or quantities adjust immediately to quality changes or the introduction of substitutes is fundamental to many elements of the methodology and its failure to hold sometimes is at the heart of many of the issues discussed in this report. We recommend that BLS identify the methodological changes required to relax this major assumption as research warrants.

### **12. The BLS should develop a number of new data collection initiatives to make some progress along these lines.**

First and foremost, BLS or a companion agency will need to collect data on detailed time use from a large sample of consumers. We would also need to extend the current health status survey to include more information on various "quality of life" issues. Progress should also be made (perhaps jointly with the BEA, which is already doing work in this area such as environmental satellite accounts) on incorporating data from victimization surveys and from various measures of the status of our physical environment into an experimental set of national satellite social-economics accounts,

accounts that value not only the market consumption basket, but also the resulting leisure and quality of life experienced by the average individual. Such accounts could also provide information on the distribution of these measures across different age and social groups. It will be difficult to integrate these into the main cost of living framework, but over time, progress on these fronts should provide useful supplementary information to policy makers and the public.

### *Suggestions for Congress*

**13. Congress should enact the legislation necessary for the Departments of Commerce and Labor to share information in the interest of improving accuracy and timeliness of economic statistics and to reduce the resources consumed in their development and production.**

**14. Congress should provide the additional resources necessary to expand the CES sample and the detail collected, to make the POPS survey more frequent, and to acquire additional commodity detail from alternative national sources, such as industry surveys and scanner data.**

While the Commission has identified some potential areas of cost savings, and it sympathizes with the recent trend to use private business sector methods to make the federal government more efficient, it notes the overwhelming trend for private businesses to be investing heavily on information technology, from tangible capital such as hardware to intangibles such as increasingly important software, to human capital.

**15. Congress should establish a permanent (rotating) independent committee or commission of experts to review progress in this area every three years or so and advise it on the appropriate interpretation of the then current statistics.**

This would be useful in its own right, but especially so to smooth the transition to a new index.

**16. Congress and the President must decide whether they wish to continue the widespread overindexing of various federal spending programs and features of the tax code. If the purpose of indexing is fully and accurately to insulate the groups receiving transfer payments and paying taxes, no more and no less, they should pass legislation adjusting indexing provisions accordingly.**

This could be done in the context of subtracting an amount partly or wholly reflecting the overindexing from the current CPI-based indexing. Alternatively, a smaller amount could be subtracted from the new revised annual index if and when it is developed and published regularly, to more closely approximate a true cost-of-living index.

We hasten to add that the indexed programs have many other features and raise many other issues beyond the narrow scope of a more accurate cost of living index. We also wish to express our view that these findings and their implications need to be digested and understood by the BLS, the Congress, the Executive Branch and the public.

#### ***To the Economists and Statisticians***

These professions should treat training in data collection, data analysis, and interpretation more seriously and give it more space and attention in the standard curriculum. There should be more emphasis on measurement and sampling issues in the training of economists and statisticians. Effort should also be put into improving the ties between professionals in government and their academic and business colleagues. The academic world needs to be cognizant of the important work done by its colleagues in government who provide them with much of the "raw material" for their subsequent analyses and show more appreciation of their efforts and understanding of the constraints under which they are laboring.

#### **IX. Conclusion**

While the CPI is the best measure currently available, it is not a true cost of living index. It suffers from a variety of conceptual and practical problems as the vehicle for

measuring changes in the cost of living. Despite important BLS updates and improvements in the Consumer Price Index, it is likely that changes in the CPI have substantially overstated the actual rate of price inflation. Moreover, revisions have not been carried out in a way that can provide an internally consistent series on the cost of living over an extended span of time. More importantly, **changes in the Consumer Price Index are likely to continue to overstate the change in the true cost of living for the foreseeable future.** This overstatement will have important unintended consequences, including overindexing government outlays and tax brackets and increasing the federal deficit and debt. If the intent of such indexing is to insulate recipients and taxpayers from changes in the cost of living, use of the Consumer Price Index has in the past, and will in the future, substantially overcompensate (on average) for changes in the true cost of living.

This report has laid out a variety of issues to be addressed in developing a more accurate measure of the cost of living. It has also presented a series of recommendations to the agency responsible for the nation's price statistics and to the elected officials who are funders, supervisors, and consumers of those statistics. We have no doubt that implementation of our recommendations would greatly improve the accuracy of the nation's price statistics. This in turn would lead to more accurate measurement of everything from cost-of-living adjustments in private contracts and public programs to information for the Federal Reserve to improved inputs to the nation's national income and product accounts. These improvements in turn would better enable citizens and policy makers alike to measure economic progress over time, among groups, and across countries.

While the commission's best estimate of the overstatement of changes in the cost of living based on changes in the consumer price index is a little over one percent, our broader point is that even small differences compound over time and matter a lot. This was evidenced in Section II when the improvement in the treatment of owner-occupied housing was introduced early in the 1980's. The same is true of the recently corrected formula bias issues which added an additional bias of about 0.24 percent per year for 1979-95. The cumulative ramifications are substantial.

While subsequent analysis, research, and economic events may result in a slight change in these estimates -- at least as likely to be up as down in our opinion -- some care in their use is warranted. While the analysis in this report represents our best judgment, this Commission did not have the substantial resources that the previous major effort to examine the nation's price statistics, the so-called Stigler Commission, had in 1961. The Stigler Commission was able to commission and produce substantial original research, while this Commission did not have the time or resources to do so. Nevertheless, this report incorporates new information from a wide variety of sources, both within the government and from outside the government. We are gratified by the tremendous outpouring of suggestions, advice, and assessment of individual issues that have arisen in the course of the committee's investigation.

The readers of this report need some time to digest and understand the results, analysis, and recommendations. This includes the BLS, the Congress, the Executive Branch, the private sector, and academe. We very much hope that careful and thoughtful consideration will be given to the findings presented in this report in the spirit in which they are offered: an attempt to provide some guidelines on how to improve the production and use of the nation's price statistics and the continuous process of improving them in a complex, dynamic economy.

---

## References

Advisory Commission To Study The Consumer Price Index. 1995. "Toward a More Accurate Measure of the Cost of Living, Interim Report," September 15, 1995.

Aizcorbe, Ana M. and Patrick C. Jackman. 1993. "The Commodity Substitution Effect in CPI Data, 1982 - 1991," *Monthly Labor Review*, Vol. 116, pp. 25 - 33 (December 1993).

Armkecht, Paul A. 1996. "Improving the Efficiency of the U.S. CPI in the Future," paper presented at NBER Productivity Workshop, April 19, 1996.

Armkecht, Paul A. and Donald Weyback. 1989. "Adjustments for Quality Change in the U.S. Consumer Price Index," *Journal Of Official Statistics*, Vol. 5, No. 2, pp. 107 - 123 (1989).

Armkecht, Paul A., Brent R. Moulton, and Kenneth J. Stewart. 1995. "Improvements

- to the Food at Home, Shelter, and Prescription Drug Indexes in the U.S. Consumer Price Index," Working Paper 263, Bureau of Labor Statistics (1995).
- Berndt, Ernst and Zvi Griliches. 1993. "Price Indexes for Microcomputers: An Exploratory Study," in *Price Measurements and Their Uses*, edited by M. Foss, M. Manser, and A. Young; NBER Studies in Income and Wealth vol. 57. Univ of Chicago Press, 1993, pp. 63 - 99.
- Berndt, Ernst R., Iain Cockburn, and Zvi Griliches. 1996. "Pharmaceutical innovations and market dynamics: Tracking effects on price indexes for anti-depressant drugs," *Brookings Papers on Economic Activity*, 1996.
- Berndt, Ernst R., Zvi Griliches, and Joshua G. Rosett. 1993. "Auditing the Producer Price Index: Micro Evidence from Prescription Pharmaceutical Preparations," *Journal of Business and Economic Statistics*, 1993 11(3), 251-264.
- Berndt, Ernst R., Zvi Griliches, and Neal Rappaport. 1995. "Econometric estimates of price indexes for personal computers in the 1990's," *Journal of Econometrics*, 68 (1995) 243-68.
- Berry, Steven, Samuel Kortum, and Ariel Pakes. 1996. "Environmental Change and Hedonic Cost Functions for Automobiles," NBER working paper 5746, September 1996, forthcoming in Proceedings of the National Academy of Sciences, 1996.
- Boskin, M. and M. Hurd. 1985. "Indexing Social Security Benefits: A Separate Price Index for the Elderly," *Public Finance Quarterly*, Volume 13, Number 4, pp. 436-449 (October 1985).
- Brown, Claire. 1994. *American Standards of Living*, Oxford, UK: Basil Blackwell Ltd., 1994.
- Carruthers, A. G., D. J. Sellwood, and P. W. Ward. 1980. "Recent Developments in the Retail Price Index," *The Statistician*, Vol 29 (1980), pp. 1-32.
- Congressional Budget Office. 1994. "Is the Growth of the CPI a Biased Measure of Changes in the Cost of Living?" *CBO Papers*, Washington. Congress of the United States, October, p. 32.
- Congressional Budget Office. 1995. *The Economic and Budget Outlook: Fiscal Years 1996-2000*, Report to the Senate and House Committees on the Budget, Washington, Congress of the United States, January 1995, Table 2-8, p. 43.

Cutler, David and Mark McClellan. 1996. "The determinants of technological change in heart attack treatment," NBER working paper #5751, Cambridge, MA, Sept. 1996.

Cutler, David M., Mark McClellan, Joseph P. Newhouse, and Dahlia Remler. 1996. "Are Medical Prices Declining?" NBER working paper 5750, September 1996.

Dahlen, J. 1994. "Sensitivity Analysis for Harmonising European Consumer Price Indices," paper presented at the International Conference on Price Indices (October 31-November 2, 1994), Ottawa, Statistics Canada.

Diewert, Erwin. 1976. "Exact and Superlative Index Numbers," *Journal Of Econometrics*, Vol. 4, No. 2, pp. 115-145 (1976).

Diewert, Erwin. 1996. "Comment on CPI Biases," *Business Economics*, Vol. 31, No. 2, pp. 30-35 (April 1996).

Diewert, W. Erwin. 1995. "Axiomatic and Economic Approaches to Elementary Price Indexes," Working Paper No. 5104, Cambridge, National Bureau of Economic Research, Inc. (May 1995).)

Division of Consumer Prices and Price Indexes. 1996a. "Changing the Hospital and Related Services Component of the Consumer Price Index," Bureau of Labor Statistics, Press Release (July 16, 1996).

Division of Consumer Prices and Price Indexes. 1996b. "Extending the Improvements in the CPI Sample Rotation Procedures and Improving the Procedures for Substitute Items," Bureau of Labor Statistics, Press Release, March 29, 1996.

Duggan, James, Robert Gillingham, and John Greenlees. 1995. "Housing Bias in the CPI and its Effect on the Budget Deficit and the Social Security Trust Fund," Office of Economic Policy, U.S. Department of the Treasury, June 30, 1995, p. 6.

Fisher, F.M. and Zvi Griliches. 1995. "Aggregate Price Indices, New Goods, and Generics," *Quarterly Journal of Econometrics* (1995).

Fisher, Irving. 1922. *The Making Of Index Numbers*, Boston, Houghton-Mifflin, 1922.

Gillingham, Robert and Walter Lane. 1982. "Changing the Treatment of Shelter Costs for Homeowners in the CPI," *Monthly Labor Review*, June 1982, p. 9.

Gillingham, Robert. 1974. "A Conceptual Framework for the Consumer Price Index," *Proceedings Of The Business And Economics Section*, American Statistical Association, pp. 246-252 (1974).

Gordon, Robert J. 1990. *The Measurement of Durable Goods Prices*. University of Chicago Press for NBER, 1990.

Gordon, Robert J. 1996. "The Sears Catalog Revisited: Apparel and Durable Goods," mimeo, Northwestern University, April 1996, revised August, 1996.

Griliches, Zvi and Iain Cockburn. 1994. "Generics and New Goods in Pharmaceutical Price Indexes," *American Economic Review*, 1994, 84(5), 1213-1232.

Griliches, Zvi and Iain Cockburn. 1996. "Generics and the Producer Price Index for Pharmaceuticals," in *Competitive Strategies in the Pharmaceutical Industry*, edited by Robert Helms, Washington, DC: AEI Press, 1996, pp. 19-34.

Hausman, Jerry A. 1966. "Valuation of New Goods under Perfect and Imperfect Competition," in Timothy Bresnahan and Robert J. Gordon, eds., *The Economics Of New Goods*, Chicago, University of Chicago Press (1966).

Hausman, Jerry. 1996. "Valuation of New Goods Under Perfect and Imperfect Competition," in T. Bresnahan and R. J. Gordon, *The Economics of New Goods*, Studies in Income and Wealth (University of Chicago Press for NBER, 1996).

Jorgenson, Dale W. and Daniel T. Slesnick. 1983. "Individual and Social Cost-of-Living Indexes," *Price Level Measurement*, W.E. Diewert and C. Montmarquette (eds.), Ottawa, Statistics Canada, pp. 241-336 (1983).

Lebow, David E., John M. Roberts, and David J. Stockton. 1994. "Monetary Policy and 'The Price Level,'" Washington, D.C., Board of Governors of the Federal Reserve System (1994).

Liegey, Paul F., Jr. 1990. "Adjusting Apparel Indexes in the Consumer Price Index for Quality Differences," in Murray F. Foss, Marilyn E. Manser, and Allan H. Young (eds.), *Price Measurements and Their Uses*. Chicago: University of Chicago Press for NBER.

\_\_\_\_\_. 1994. "Apparel Price Indexes: Effects of Hedonic Adjustment," *Monthly Labor Review* (May), pp. 38-45.

Moulton, Brent R. 1996. "Estimation of Elementary Indexes of the Consumer Price Index," Bureau of Labor Statistics (May 1996.)

Moulton, Brent R. and Karin E. Smedley. 1955. "A Comparison of Estimator for Elementary Aggregates of the CPI," Bureau of Labor Statistics (June 1955).

- Moulton, Brent R. and Karin E. Smedley. 1995. "A Comparison of Estimators for Elementary Aggregates of the CPI," Bureau of Labor Statistics, (June 1995).
- Nordhaus, William D. 1996. "Do Real-Output and Real-Wage Measures Capture Reality? The History of Light Suggests Not," in T. Bresnahan and R. J. Gordon, *The Economics of New Goods*, Studies in Income and Wealth (University of Chicago Press for NBER, 1996).
- O'Neill, June. 1995. "Prepared Statement," *Consumer Price Index*, Hearings before the Committee on Finance, U.S. Senate, 104th Congress, First Session, Washington, U.S.G.P.O., Table 1, p. 146.
- Pakes, Ariel, Steven Berry, and James A. Levinsohn. 1993. "Applications and Limitations of Some Recent Advances in Empirical Industrial Organization: Price Indexes and the Analysis of Environmental Change," *American Economic Review*, Vol. 83, No. 2, pp. 241-246 (May 1993).
- Randolph, William C. 1988. "Housing Depreciation and Aging Bias in the Consumer Price Index," *Journal of Business and Economic Statistics* (July 1988), vol. 6, no. 3, pp. 359-71.
- Reinsdorf, Marshall B. 1993. "The Effect of Output Price Differentials on the U.S. Consumer Price Index," in Murray F. Foss, Marilyn E. Manser, and Allan H. Young, eds., *Price Measurements And Their Uses*, Studies in Income and Wealth vol. 57, Chicago, University of Chicago Press for NBER.
- Repetto, Robert, et. al. 1990. *Has Environmental Protection Really Reduced Productivity Growth?* World Resources Institute, 1990.
- Schultz, B. J. 1994. "Choice of Price Index Formula at the Micro- Aggregation Level: The Canadian Empirical Evidence," paper presented at the International Conference on Price Indices (October 31-November 2, 1994), Ottawa, Statistics Canada.
- Shapiro, M. D. and D. W. Wilcox. 1996a. "Alternative Strategies for Aggregating Prices in the CPI," paper prepared for the conference on Measuring Inflation and Real Growth, Federal Reserve Bank of St. Louis, October, 1996.
- Shapiro, Matthew D. and David W. Wilcox. 1996b. "Mismeasurement in the Consumer Price Index: An Evaluation," Working Paper No. 5590, Cambridge, National Bureau of Economic Research, Inc. (May 1996).

- Shapiro, Matthew D., and David W. Wilcox. 1996c. "Mismeasurement in the Consumer Price Index: An Evaluation." *NBER Macroeconomics Annual 1996*. (SAME AS ABOVE?)
- Slifman, L., and C. Corrado. 1996. "Decomposition of Productivity and Unit Costs," mimeo, Federal Reserve Board of Governors, November 1996.
- Statistical Abstract of the United States*. 1978.
- Statistical Abstract of the United States*. 1990.
- Statistical Abstract of the United States*. 1995.
- The U. S. Consumer Electronics Industry In Review 1995*, Electronic Industries Association, p. 13.
- Tornqvist, Leo. 1936. "The Bank of Finland's Consumption Price Index," *Bank Of Finland Monthly Bulletin*, Vol. 10, pp. 1-8 (1936).
- Trajtenberg, Manuel. 1990. "Product Innovations, Price Indices, and the (Mis)Measurement of Economic Performance," NBER Working Paper no. 3261, February 1990.
- Triplett, Jack. 1990. "The Theory of Industrial and Occupational Classifications and Related Phenomena," *Proceedings of the Bureau of the Census 1990 Annual Research Conference*, August 1990.
- U. S. Census Bureau. 1993. *American Housing Survey, 1993*.
- U. S. Census of Housing, 1970*.
- Woolford, K. 1994. "A Pragmatic Approach to the Selection of Appropriate Index Formulae," paper presented at the International Conference on Price Indices (October 31-November 2, 1994), Ottawa, Statistics Canada.

## **Members of the Advisory Commission To Study The Consumer Price Index**

### **Michael J. Boskin, Ph.D., Chairman**

Tully M. Friedman Professor of Economics  
 & Senior Fellow, Hoover Institution  
 Stanford University  
 Stanford, California

**Ellen R. Dulberger, Ph.D.**

Director of Marketing Strategy  
IBM Personal Computer Company  
Somers, New York

**Robert J. Gordon, Ph.D.**

Chairman, Department of Economics  
& Stanley G. Harris Professor in the Social Sciences  
Northwestern University  
Evanston, Illinois

**Zvi Griliches, Ph.D.**

Paul M. Warburg Professor of Economics  
Harvard University  
Cambridge, Massachusetts

**Dale Jorgenson, Ph.D.**

Chairman, Department of Economics  
& Frederic Eaton Abbe Professor of Economics  
Harvard University  
Cambridge, Massachusetts

---

**FOOTNOTES**

1 [BACK](#) We would like to thank the staffs of the Bureau of Labor Statistics, Congressional Congressional Budget Office, Senate Finance Committee and other individuals in academic and the private sector too numerous to mention hem for valuable assistance and advise during the Commission's work.

2 [BACK](#) The two most commonly used are the CPI-U and CPI-W. The former is for all urban consumers, roughly 80% of the population; the latter is for urban wage and clerical workers, about 32% of the population. Note that ft expenditure shares may be quite different than the average for any particular household, and also on average for subgroups of the population. Also, the prices paid for some products may differ for some households from the prices actually sampled. In principle, if not practice, a separate cost of living index could be developed for each and every household based on their actual consumption basket and prices paid. The overall index is used to approximate this with the data reflecting representative consumers. Whether this is itself sufficiently misleading as to warrant separate price indexes

for different population subgroups is discussed below.

3 [BACK](#) See Congressional Budget Office (1994).

4 [BACK](#) See Gillingham and Lane (1982).

5 [BACK](#) See Duggan, Gillingham and Greenlees, (1995).

6 [BACK](#) See Congressional Budget Office (1995).

7 [BACK](#) See O'Neill (1995). These CBO budget estimates are relative to CBO's January 1995 baseline and do not include the small adjust assumed in the out-years of the budget resolution.

8 [BACK](#) These estimates are relative to the CBO's May 1996 Baseline. See CBO (1996).

9 [BACK](#) Chapter 19 of BLS Handbook of Methods. We are especially grateful to John Greenlees and Brent Moulton of BLS for clarifying several of these issues for the Commission.

10 [BACK](#) Prices are actually collected in 88 locations, called primary sampling units, or PSUs. In eight PSUs (the five largest urban areas), prices are collected for all items every month. In other areas, prices are collected monthly for food, fuels and a few other items, and bimonthly for all other items. Of the 44 areas which go into the index every month, 32 are self-representing because of their size and 12 are composites constructed from 56 PSUs which provide representation for smaller and mid-size cities across the country.

11 [BACK](#) The housing component properly captures the multiple period consumption accrued subsequent to the purchase of a house. This same approach is warranted but not used in measuring durable goods such as automobiles and refrigerators.

12 [BACK](#) 184 of the item groups are actually priced. The other 23 strata account for less than 2 percent of the weight of the overall index. Price indexes for these groups are moved with changes in the indexes for the item groups which are priced.

13 [BACK](#) This example is Current and real, provided by BLS. The names of the brand and store are withheld to adhere to confidentiality requirements.

14 [BACK](#) There are other reasons that may result in the disappearance of a specific item from a specific outlet. When the reason is loss of competitive market share, the BLS replacement procedures are likely to result in upward bias.

15 [BACK](#) This procedure is based on the assumption that the marketplace adjusts fully and instantaneously to price differences among competing products. This can happen in two ways: prices of substitutes change immediately to make them equal (quality-adjusted), or quantities of what would be higher priced products fall to zero making them disappear from the market.

16 [BACK](#) The market basket has been updated once per decade historically and introduced with a several year lag. For current data, it represents an average derived from surveys for 1982-1984. The next revision of the index is scheduled for introduction in 1998 at which time the base period will be updated to 1993-1995.

17 [BACK](#) Only seasonal factors are revised historically.

18 [BACK](#) Sometimes a new product is introduced with sample rotation. An example might be a new variety of apples. 19 [BACK](#) The ability of the CPI to fully capture price impacts such as these depends on the degree to which the classification structure is consumption-based. That is, items which are the closest substitutes for each other in terms of how they are used, must be in the same item group, the lowest level at which indexes are constructed. The item structure is updated with the decadal revisions of the CPI. The new item structure which will be introduced with the 1998 revision will make some improvements toward placing close substitutes together. Much more is needed.

20 [BACK](#) Another way to state this is that the elasticity of substitution among items within the lowest grouping, say types of apples, is assumed to be zero.

21 [BACK](#) Each method has come to be named for its inventor. See below.

22 [BACK](#) This index is called the Laspeyres index.

23 [BACK](#) This index is called the Paasche index.

24 [BACK](#) See Fisher (1922).

25 [BACK](#) See Diewert (1976).

26 [BACK](#) See Tornqvist (1936).

27 [BACK](#) See Moulton and Smedley (1955).

28 [BACK](#) See Diewert (1995).

29 [BACK](#) See Shapiro and Wilcox (1996b).

30 [BACK](#) See Carruthers, Sellwood, and Ward (1980), Schultz (1994), Dahlen (1994), and Woolford (1994).

31 [BACK](#) For detailed discussion of Lower Level Substitution Bias, see: Moulton (1996).

32 [BACK](#) See Moulton and Smedley (1995).

33 [BACK](#) This undoubtedly arose both to save overtime pay and to accommodate retailers who prefer BLS price-takers to be in their establishments when they are less crowded.

34 [BACK](#) See Jorgensen and Slesnick (1983).

35 [BACK](#) See Fisher and Griliches (1995).

36 [BACK](#) There is no presumption that the magnitude of upward quality bias has declined over

time. One consideration is that the growing importance of such hard-to-measure categories as consumer electronics and medical services may have increased the significance of quality change bias in the past decade. Another problem is suggested by a "thought experiment" recently conducted by Nordhaus (1996), who extrapolated backward substantial upward bias in the CPI over a long period of 190 years and arrived at implausibly low estimates of the standard of living of the average U. S. citizen in the year 1800. The implausibility of continuous upward bias in earlier decades at the rate suggested for recent decades in this report implies that in some earlier era the upward bias in the CPI was substantially less. This, of course, is natural. Long ago more was more important than better, e.g., enough to eat was more important than variety. As incomes rise beyond some point it is natural to expect increased demand for quality in many goods and services. We return to this issue below in our discussion of housing prices.

37 [BACK](#) Consumer Reports since 1959 has based its quotation for the average price of a particular model on an average of prices obtained in a market survey, typically of 13-20 price quotations obtained across the country. Thus a shift to discount stores would show up in the Consumer Reports price quotations and account for part of the difference between Gordon's indexes and the CPI index for the same category.

38 [BACK](#) As a further example of the difficulty of disentangling quality change from other sources of bias, one important fact to be assessed is the relatively large recent difference between the growth rates of the CPI and deflator for Personal Consumption Expenditures (hereafter PCE deflator). The PCE deflator is part of the National Income and Product Accounts, which is the responsibility of the Bureau of Economic Analysis, a division of the Department of Commerce. Over the 12 months ending in August, 1996, the CPI rose at 2.9 percent per year, while in the four quarters ending in 1996: Q3, the implicit PCE deflator rose at 2.0 percent per year. An alternative PCE deflator excluding expenditures on personal computers and medical care increased at 2.7 percent per year, suggesting that much of the difference between the PCE deflator and CPI over this period may involve their differing treatments of computers and medical care. This calculation was made by Bob Arnold of the CBO by subtracting nominal and real expenditures on medical care and consumer purchases of computers from total personal consumption expenditures. But that is not just an issue of differing treatment of quality change, but also substitution bias, since consumer expenditures on both personal computers and medical care have increased greatly the 1982-84 base year of the CPI.

39 [BACK](#) The number of cellular phone subscribers is taken from Business Week, December

2, 1996, p. 104.

40 [BACK](#) All annual growth rates calculated in this report are logarithmic.

41 [BACK](#) Brown's (1994) compilation of annual rent paid by a working class household yields \$5160 for 1988 (Table 7.9a) and \$1803 for 1973 (Table 6.9a), for a ratio of 2.86. The ratio of the CPI unadjusted rent index for the two years is 2.43.

42 [BACK](#) All citations regarding housing quality in this section to the U. S. Census of Housing, 1970 and the U. S. Census Bureau, American Housing Survey, 1993.

43 [BACK](#) Data from Statistical Abstract of the United States. 1978 issue, Tables 1371 and 1373. 1995 issue, Tables 1214 and 1230.

44 [BACK](#) The 1918 and 1950 data are for annual rental expenditures of working class households, from Brown (1994), Tables 3.6A and 5. 10.

45 [BACK](#) Brown (1994), p. 63.

46 [BACK](#) This estimate of 1.0 percent per year multiplies an estimate of 1.5 percent per year for telephone service by its weight in this category of roughly two-thirds.

47 [BACK](#) The U. S. Consumer Electronics Industry In Review 1995, Electronic Industries Association, p. 13.

48 [BACK](#) Data provided by the BEA. Audio-video equipment includes VCRs, camcorders, and videotapes but excludes 'audio media,' i.e., cassettes and CDs. By 1996:Q3 computers accounted for 0.43 percent of nominal PCE and 1.34 percent of real PCE in 1992 dollars.

49 [BACK](#) While the CPI incorporates a matched-model index that records a decline in computer prices of 10-20 percent per year, this is applied to a negligible weight (based currently on 1982-84 weights) and so has no practical importance for assessing the magnitude of the quality change bias in the personal computer category. Berndt, Griliches, and Rappaport (1995) have estimated an annual difference between 'matched model' indexes for personal computers (of the type used by the CPI) and hedonic indexes (used by the BEA) of roughly 15 percent per year.

50 [BACK](#) Recent changes in the CPI treatment of apparel are discussed in Liegey (1990, 1994).

51 [BACK](#) A total of 1,769 matched price comparisons were made for 1914-47 and 4,640 for 1965-93.

52 [BACK](#) Recall from footnote 49 that the CPI treatment of apparel changed in 1985.

53 [BACK](#) The 1967-93 annual rates of growth for the 'raw' price of new cars 6.65 percent the safety-pollution adjustment -2.03 percent, the fuel economy adjustment -1.10 percent (totalling 3.52 percent), and the new car CPI 4.42 percent.

54 [BACK](#) The growth for trucks are similar. Data for 1970 and 1983 come from Statistical Abstract, 1990, Table 1027, and for 1993 from Statistical Abstract, 1995, Table 1025.

55 [BACK](#) In 1987 the BLS began to adjust for quality changes in used cars in a parallel manner to its adjustment for new cars.

56 [BACK](#) The 1982 Census of Retailing reports that the number of gas pumps in the United States declined 28 percent between 1972 and 1982, and that the fraction of self-service pumps (which was not reported in 1972) had reached 54 percent by 1982.

57 [BACK](#) The BLS is moving to price and reprice a hospital bill every month, a potential improvement, although not something which per se gets at quality change.

58 [BACK](#) In addition to the comprehensive bias estimates presented in our Interim Report, estimates have been made by Erwin Diewert (1996), Shapiro and Wilcox (1996), and others. An especially valuable earlier survey was presented by Lebow, Roberts, and Stockton (1994).

59 [BACK](#) Aizcorbe and Jackman (1993) have estimated Upper Level Substitution Bias, but these estimates have been updated and revised by BLS.

60 [BACK](#) For details, see Gillingham (1974).

61 [BACK](#) See Reinsdorf (1993), pp. 227-254.

62 [BACK](#) The results of BLS research on Lower Level Substitution Bias are summarized in Moulton (1996).

63 [BACK](#) See Armknecht, Moulton, and Stewart (1995).

64 [BACK](#) See Division of Consumer Prices and Price Indexes (1996b).

65 [BACK](#) See Division of Consumer Prices and Price Indexes (1996a).

66 [BACK](#) See Armknecht and Weyback (1989).

67 [BACK](#) See Hausman (1996), Griliches and Cockburn (1994), and Pakes, Berry and Levinsohn (1993).

68 [BACK](#) Although the BLS is continuously attempting to improve the price measures, for example moving to price and reprice hospital bills.

69 [BACK](#) See Boskin and Hurd (1985); Jorgensen and Slesnick (1983). However, very preliminary unpublished work suggests that for the period 1982-91 the larger fraction of expenditures on out-of-pocket healthcare by the elderly combined with the more rapid rise in healthcare prices than overall prices for this period might lead to a slightly faster rise in a price index for the elderly. The rate of healthcare price inflation has slowed substantially of late, so it is unlikely this result will be reproduced in the mid-1990s. Also, as noted above in Section V, healthcare inflation is seriously overstated because of the substantial uncounted

quality change.

70 [BACK](#) Seniors get special discounts, for example, and their geographic distribution, and other factors might cause the prices they pay to differ from those recorded in the CPI.

71 [BACK](#) Detailed case studies in Repetto et. al. (1996) demonstrate substantial understatement of output and productivity growth, and hence and overstatement of price growth, in the electricity, pulp-paper, and agricultural industries.

72 [BACK](#) All data in the paragraph come from the Statistical Abstract, 1995, Table 317.

73 [BACK](#) Data in this paragraph come from the Statistical Abstract, 1995, Tables 114, 124, 136, and 141.

74 [BACK](#) A parallel trailing chained-link Laspeyres formula, with revisions producing a comparable Fisher Ideal Index with a 4-year lag might be useful as well. Even so, we will still recommend the move to geometric means at the elementary aggregates level.

75 [BACK](#) What needs more rapid changing are the commodity strata weights and the assortment of items priced. City weights could be held constant and changed much less frequently. Here decades would do. See Shapiro and Wilcox (1996a) for additional discussion of the mechanics of such revisions.