

## 국제 원유 가격들의 확산과정 모형 추정\*

최승문\*\*

### 요 약

본 연구의 목적은 국제 원유 가격의 동태적 변화를 잘 설명할 수 있는 확산과정 (diffusion process) 모형을 실증분석을 통해 찾는 것이다. 이를 위해 유가의 변화에 대해 얘기할 때 대표적으로 언급되는 두바이유, 브렌트유 그리고 WTI유의 일별 자료를 이용해 선행 연구에서 이용된 대부분의 확산과정 모형들을 포함하는 GD-GV 모형과 여기에 포함되는 GD-CEV 모형, CKLS 모형, CIR 모형 그리고 Vasicek 모형을 추정했다. 최우추정법을 이용했는데 이를 위해 필요한 로그전이확률밀도 함수는 Ait-Sahalia (2008)가 개발한 방법을 이용해 근사적으로 구했다. 세 종류의 유가 모두에서 비슷한 추정 결과들을 얻었다. CKLS 모형이 유가의 움직임을 가장 잘 설명한다는 결과를 얻었다. 그리고 추세 함수 보다는 변동성 함수가 유가의 움직임에 더 중요한 역할을 한다는 사실을 알았다. 유가가 낮을 때는 평균값으로 회귀한다는 증거를 찾을 수 있었지만 중간 값 이후에서는 그러한 증거를 찾을 수 없었다. 유가에 대한 변동성의 탄력성은 약 0.75로 추정됐다.

주요 단어 : 원유 가격, 확산과정, 최우추정법

경제학문헌목록 주제분류 : Q41, C52

\* 본 논문의 개선을 위해 유익한 조언을 해 주신 익명의 심사위원님들 세 분과 에너지경제 연구의 편집장님께 깊이 감사드립니다. 이 논문은 2020년도 서울시립대학교 교내학술연구비에 의하여 지원되었음.

\*\* 서울시립대학교 경제학부 교수. schoi22@uos.ac.kr

## I. 서 론

Black and Scholes (1973)와 Merton (1973)이 주식 가격이 연속시간 확산 과정 (diffusion process)의 하나인 기하 브라운운동 (Geometric Brownian motion) 모형을 따른다는 가정 하에 주식에 대한 유럽식 옵션의 가격을 구체적인 식으로 구한 이후에 확산과정 모형은 특히 금융 분야에서 다양한 이론을 개발하는데 많이 활용됐다. 주가, 환율, 이자율과 같은 금융 변수들의 움직임을 확산과정으로 모형화하면 이들 기초자산의 선물, 옵션 그리고 다양한 파생상품들의 가치를 계산할 때 편리하다. 이러한 금융 자산들뿐만 아니라 옥수수, 밀, 커피와 같은 농산물, 소, 돼지와 같은 가축, 금, 구리, 알루미늄과 같은 금속, 원유, 천연가스, 전기와 같은 에너지관련 생산물과 같이 다양한 현물들 그리고 이들과 관련된 여러 파생 상품들이 거래되고 있다. 이러한 현물들의 가격의 움직임은 금융 변수들과 비슷한 성질들을 가지고 있어 확산과정으로 모형화해 이들과 관련한 파생상품들의 가격을 쉽게 계산할 수 있다.

원유와 관련한 파생상품들의 가격을 정확히 계산하기 위해 또는 원유 개발 사업의 가치를 올바르게 평가하기 위해 원유 가격의 움직임을 잘 설명할 수 있는 확산과정 모형을 이용하는 것이 중요하다. 이 논문에서는 원유 가격의 움직임을 잘 설명하는 확산과정 모형이 무엇인지 원유가 자료가 말을 할 수 있도록 국제 원유가의 동향을 파악할 때 대표적으로 가장 많이 언급되는 중동의 두바이 (Dubai)유, 영국 북해의 브렌트 (Brent)유 그리고 미국 서부 텍사스 지역의 WTI유의 일별자료를 이용해 다섯 개의 단일변수 시간균질 확산과정 모형들을 추정하고 비교한다. 이들은 Ait-Sahalia (1996)가 제안한 모형에 Choi (2009)가 추세함수에 3차 항을 추가한 일반적 추세함수 (general drift: GD)와 일반적 변동성 (general volatility: GV) 함수를 갖는 GD-GV 모형 그

리고 GD-GV 모형과 추세함수는 같지만 분산의 유가에 대한 탄력성이 상수인 CEV (constant elasticity of variance) 변동성 함수를 갖는 GD-CEV 모형, Chan, Karolyi, Longstaff, and Sanders (1992)이 제안한 CKLS 모형, Cox, Ingersoll, and Ross (1985)가 제안한 CIR 모형, 그리고 Vasicek (1977)이 제안한 모형이다. GD-GV 모형은 나머지 네 개의 모형들뿐만 아니라 선행 연구들에서 여러 금융 또는 현물 변수들의 모형을 위해 이용된 대부분의 확산과정 모형들을 포함하는 매우 일반적인 것이다.

확산과정을 이용한 유가와 관련한 선행연구들을 보면 Brennan and Schwartz (1985)는 광산이나 다른 천연 자원 개발 사업의 가치를 평가하기 위해 관련된 현물의 가격이 기하 브라운운동을 따른다는 가정을 했다. 해상 석유 개발권의 가치를 계산하기 위해 Paddock et. al. (1988)은 원유가가 확산과정을 따른다고 가정한다. Dias and Nunes (2011)은 현물가격의 특성을 잘 설명할 수 있도록 Black-Scholes-Merton 모형에서 변동성 함수를 CEV로 바꾼 모형으로 미국식 옵션 가격을 구했다. Dixit and Pindyck (1994), Cortazar and Schwartz (1997) 그리고 Smith and McCardle (1999)은 원유 가격이 장기적인 평균 수준으로 회귀한다는 성질을 포착할 수 있는 모형도 고려했다. 원유의 가격에 따라 결정되는 금융 그리고 실물 자산의 가치를 계산하기 위해 Gibson and Schwartz (1990)는 원유 가격이 기하 브라운운동을 따르고 여기에 물리적으로 원유를 가지고 있음으로써 얻는 이익을 나타내는 편의 수익 (convenience yield)의 확산과정을 추가해 2요인 확산과정 모형을 개발했다. Schwartz and Smith (2000)은 유가가 0으로 회귀하는 성질을 가진 단기 이탈 요인과 유가의 균형가격 수준을 결정하는 요인 두 개의 합으로 결정되는 2요인 확산과정 모형을 제안했다. Dias and Rocha (1999)는 장기적인 평균수준으로 회귀하는 기하 브라운운동에 점프를 추가한 모형을 이용해 유가를 모형화해 원유 개발사업의 가치를 평가하고 기존의 연구결과들과 비교했다. 현물과 관련한 파생 상품들의 가격을 계산하기 위해 4요인 점프 확산과정모형을 제안한 논문은 Yan (2002)이다. 다른 한 편 Postali and Picchetti (2006)는 국

제 유가 자료의 확률적 성질들을 분석해 약 30년 정도 걸리는 유전 개발사업의 가치를 평가하기 위해 비교적 단순한 기하 브라운운동 모형도 적절하다고 주장한다. Liu et. al. (2020)은 원유 가격의 변동성에 국면전환이 있다는 증거를 제시했고 Bernabe et. al. (2004)는 원유가의 움직임을 설명하기 위한 모형으로 평균회귀 Vasicek 모형을 국면전환 (regime-switching)과 결합할 것을 제안한다. Kyriakou et. al. (2016)은 WTI유 선물 옵션 가격을 잘 설명하는 모형을 찾기 위해 Heston (1993) 모형과 여기에 점프를 추가한 Bates (1996) 모형을 추정했다. Schwartz (1997)은 원유를 포함한 현물 가격들을 이용해 이들에서 강한 평균 회귀 성질이 있음을 밝혔다.

원유 가격을 확산과정 모형에 적용한 실증분석 연구로 윤원철 (2005)은 Nowman and Wang (2001)이 다섯 종류의 금속 현물 가격들에 대해 확산과정 모형들을 추정한 것과 같은 방식으로 두바이유, 브렌트유, WTI유, 국내 도입 원유와 국내 등유 가격들의 월별 평균 가격자료를 실증분석 했다. 윤원철, 박호정 (2005)에서는 월별 브렌트유 가격 자료를 윤원철 (2005)와 같은 방식으로 추정하고 추정 결과를 바탕으로 유럽식 콜옵션의 가치를 계산하고 비교했다. 확산과정 모형을 추정하는데 고빈도 자료가 더 적절하지만 위의 선행 연구들에서는 일별 자료가 있음에도 불구하고 월별 자료를 이용하고 연속시간 모형을 이산화해 추정했기 때문에 추정 결과에 편의 (bias)가 생긴다 (Lo (1988)). Askari and Krichene (2008)은 유가 자료를 이용해 Merton (1976)의 점프 확산과정을 추정했다. Larsson and Nossman (2011)은 유가의 어파인 (affine) 점프 확률변동성 모형들을 추정해 유가의 변동성이 확률적이며 점프가 있다는 증거를 제시한다. 정세록 (2016)에 원유가와 관련된 선행 연구들이 잘 정리돼있다.

유가의 움직임을 설명하기 위해 비교적 단순한 기하브라운 운동 모형부터 점프를 포함한 다요인 모형까지 여러 종류의 확산과정 모형들이 이론 개발과 실증 분석을 위해 제안됐다. 그런데 단일 변수 모형을 이용한 연구에서도 본 논문에서 고려하는 것과 같이 일반적인 모형을 이용한 것이 아니거나 사용된

추정방법이 확산과정 모형을 추정하기에 적절하지 않다. 이 연구에서는 아주 일반적인 확산과정 모형과 이를 추정하기 위해 개발된 추정방법을 이용해 원유가들의 움직임을 설명하는 가장 적합한 확산과정 모형을 찾는다는데 의의가 있다.

우리가 사용한 추정 방법은 가장 효율적인 최우추정 (maximum likelihood estimation)법이다. 이를 위해서는 확산과정의 전이확률밀도 (transition probability density) 함수를 알아야 하는데 Vasicek 모형이나 CIR 모형과 같이 몇 개의 예외적인 경우를 제외하고는 대부분의 확산과정의 전이확률밀도 함수는 알려져 있지 않다. 그래서 Ait-Sahalia (2002)는 허마이트 (Hermite) 급수 전개를 이용해 단일변수 시간균질 (time-homogeneous) 확산과정의 전이확률밀도 함수를 구체적인 식으로 근사시키는 방법과 이론을 개발했다. 이 방법은 Ait-Sahalia (2008)에서 다변수 시간균질 확산과정으로 확장됐고, Choi (2013, 2015a)는 다변수 시간비균질 (time-inhomogeneous) 확산과정으로, Yu (2007)과 Choi (2019)은 각각 다변수 시간균질과 시간비균질 점프 확산과정으로 일반화했다. 이 외에도 이들 연구를 이용한 실증 분석 연구 등 이와 관련한 여러 연구결과들이 있는데 Choi (2020)에 좀 더 자세히 정리되어 있다. 이 논문에서는 최우추정법을 이용하기 때문에 우도함수가 필요하므로 활용하기에 편한 Ait-Sahalia (2008)가 제안한 방법을 이용해 확산과정들의 로그우도함수를 근사시켰다.

실증분석 결과 흥미롭게도 GD-GV 모형과 GD-CEV 모형이 CKLS 모형을 포함하는 더 일반적인 모형들이지만 본 연구에서 활용한 모든 원유가들에 대해 CKLS 모형이 가장 선호되는 것을 확인할 수 있었다. 추정 값에는 약간의 차이가 있었지만 추정 결과는 전반적으로 아주 비슷했다. 먼저 정보기준 AIC와 BIC 모두에 대해 CKLS 모형이 가장 작은 값을 가져 다른 네 가지 모형들보다 더 우월함을 알 수 있었다. 그리고 CKLS 모형을 포함하는 GD-GV 모형과 GD-CEV 모형에 추가된 모수들은 통계적으로 유의하지 않았다. 추세함수 보다는 변동성 함수와 관련된 모수들이 통계적으로 유의한 것을 봤을 때 변동성 함수가 유가의 움직임을 설명하는데 더 중요한 역할을 한다고 볼

수 있다. 유가가 낮을 때는 유가가 장기적인 평균 수준으로 회귀하는 성질을 확인할 수 있었으나 중간 이상의 유가에서는 그렇다는 통계적 증거를 찾을 수 없었다. 변동성 함수는 통계적으로 매우 유의하게 추정되었고 유가가 늘어날수록 변동성도 증가한다. 변동성의 유가에 대한 탄력성은 약 0.75로 유가가 1% 늘어나면 변동성은 이보다 작게 0.75%정도 증가하는 결과를 얻었다.

다음 장에서는 이 연구에서 활용한 세 종류의 유가들에 대해 좀 더 자세히 살펴볼 것이다. 3장에서는 실증분석을 위해 채택한 확산과정 모형들에 대한 설명과 추정 방법에 대해 논의하고 이어진 장에서는 추정 결과와 이에 대한 해석과 논의를 한다. 마지막으로 5장에서 결론과 추가적인 연구에 대해 논의한다.

## II. 원유 가격 자료

여러 나라들이 원유를 생산하고 국제적으로 많은 원유 제품들이 거래되고 있는데 그 중에서 국제 유가의 움직임에 대한 지표로서 대표적으로 가장 많이 언급되는 것들이 중동의 두바이 (Dubai)유, 영국 북해의 브렌트 (Brent)유 그리고 미국 서부 텍사스 지역의 WTI유의 원유 가격들이다. 본 연구에서는 한국석유공사의 Petronet에서 구한 이들 원유 현물들의 가격을 이용해 여러 확산과정 모형들을 추정했다. 모형 추정 당시 Petronet에서 구할 수 있는 모든 자료를 활용했는데, 두바이유는 1996년 1월 2일부터 2020년 2월 21일까지 그리고 브렌트유와 WTI유는 1998년 1월 2일부터 2020년 2월 21일까지의 일별 자료들을 이용했다.

〈표 1〉 국제 원유 가격들의 기초 통계량

(단위: 미국달러/배럴)

	두바이 원유 가격	브렌트 원유 가격	WTI 원유 가격
기간	1996.1.2 ~ 2020.2.21	1998.1.2 ~ 2020.2.21	1998.1.2 ~ 2020.2.21
관측치 수	6103	5652	5554
최소	9.45	9.64	10.72
최대	140.70	146.08	145.29
평균	54.84	59.89	57.71
표준편차	32.32	30.78	29.03
첨도	0.49	0.34	0.41
왜도	-0.90	-0.82	-0.78

주: 1996년 1월 2일부터 2020년 2월 21일 까지 일별 두바이 (Dubai) 원유 가격과 1998년 1월 2일부터 2020년 2월 21일 까지 일별 브렌트 (Brent)와 WTI 원유 가격의 기초 통계량들이 제시되어 있다. 이들 원유 가격의 단위는 배럴당 미국 달러다. 여기에서 왜도는 정규분포의 왜도 값 3을 뺀 초과 왜도 값이다.

〈표 1〉에는 이들 원유가격들의 기초 통계량 값들이 정리되어 있다. 주말과 각 원유가 거래되는 지역의 공휴일에는 원유 거래가 없어 이들 날에 해당하는 자료는 없기 때문에 브렌트유와 WTI유의 자료 기간이 같지만 관측치의 수는 다르다. 배럴당 미국 달러로 인용되는 이들 원유가의 구체적인 기초 통계량 값들에는 약간의 차이가 있지만 전반적으로 비슷한 것을 알 수 있다. 원유를 구성하는 여러 화학적인 물질들의 성질에 따라 원유의 가치를 다르게 평가하는데 보통 브렌트유는 WTI보다 배럴당 약 2달러 높고 두바이유 보다 는 약 5달러에서 6달러 정도 높게 책정되는 것으로 알려져 있는데 이들 원유가의 평균 가격을 보면 이를 확인할 수 있다. 원유가의 최솟값은 큰 차이가 없으며 최댓값은 브렌트유와 WTI유가 비슷하고 이들이 두바이유 보다 약 5달러 정도 높다. 표준편차는 두바이유가 가장 크고 브렌트유와 WTI유가 비슷하다. 첨도는 모두 양의 값을 가져 정규분포와 비교해 가격 분포의 오른쪽 꼬리가 더 긴 것을 알 수 있다. 왜도는 모두 정규 분포보다 작아 정규분포 보

다 극단적인 양쪽 꼬리의 값들이 덜 관측되는 것을 알 수 있다. 이들 두 값은 세 원유가들의 분포가 모두 정규분포와는 다르다는 것을 보여준다.

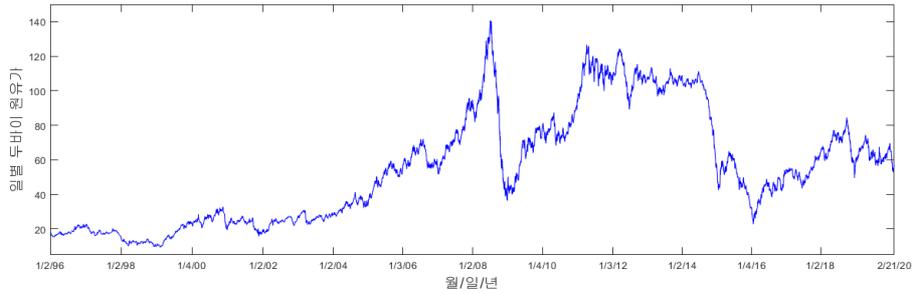
[그림 1. 가]는 1996년 1월 2일부터 2020년 2월 21일까지의 두바이유 그리고 [그림 2. 가]와 [그림 3. 가]에는 각각 1998년 1월 2일부터 2020년 2월 21일까지의 브렌트유와 WTI유의 일별 시계열 자료를 그래프로 그렸다. 이 그림들을 보면 같은 기간 동안에 모든 유가들이 값의 차이는 약간 있지만 전반적인 움직임들은 매우 비슷한 것을 확인할 수 있다<sup>1)</sup>.

이들 그림에서 원유 산업의 상황, 주요 산유국들의 상황과 국제 정세 그리고 세계 경제의 상황 등과 관련한 여러 사건들에 따라 원유가가 변해온 것을 알 수 있다. 눈에 띄는 것은 2004년부터 눈에 띄게 오르던 원유가가 2008년 세계금융위기 직전부터 떨어지기 시작하더니 본격적인 금융위기가 닥치면서 급격히 하락한 것이다. 이후 다시 오르기 시작한 원유가는 여러 나라들의 경제 문제로 인한 수요의 감소, 셰일 오일 생산, 유가 하락에 대한 OPEC의 무대응 등 여러 사건들이 결합되어 2014년 말에 2009년의 최저가 보다 더 낮은 값으로 떨어졌다.

유가의 일별 변화가 전날의 유가의 수준과 관련이 있는지 여부를 알아보기 위해 각 시점에서의 유가에 대해 다음 날 유가의 변화가 얼마가 되는지를 보여주는 산포도를 두바이유, 브렌트유 그리고 WTI유에 대해 각각 [그림 1. 나], [그림 2. 나] 그리고 [그림 3. 나]에 그렸다. 유가의 장기적인 평균값이 있어 이 값으로 회귀하는 경향이 있다면 유가의 수준이 이 평균 수준 보다 더 낮으면 다음날 유가가 증가하고 유가가 평균 수준보다 높으면 다음날 유가가 더 크게 감소해 산포도의 왼쪽 끝으로 갈수록 양의 일별 변화를 오른쪽 끝으로 갈수록 음의 일별 변화를 나타내는 점들을 더 많이 볼 수 있을 것을 예상할 수 있다. 그런데 육안으로 봤을 때 세 유가들 모두에서 유가가 높을수록

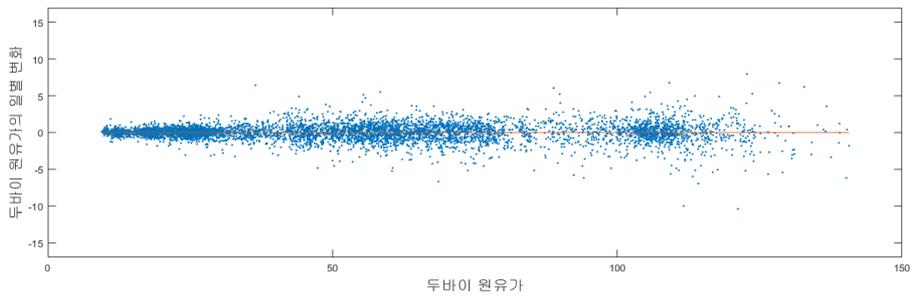
1) 본 연구에서 이용한 주요 국제 원유 가격들은 일별 자료로 이들이 거래되는 지역의 다른 공휴일 등으로 인해 다른 평일에 관측 자료가 없는 경우들이 있었기 때문에 그리고 기간을 같이 조정하더라도 관측 자료의 수가 다르기 때문에 이들 사이의 상관계수를 구할 수는 없었다.

[그림 1. 가] 두바이유의 일별 가격



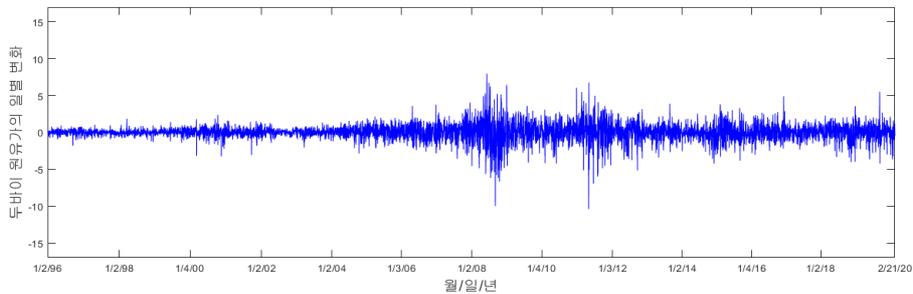
주: [그림 1. 가]에는 1996년 1월 2일부터 2020년 2월 21일 까지 배럴당 미국 달러로 표시된 두바이유 가격의 일별 시계열 자료가 그려져 있다.

[그림 1. 나] 두바이유 가격의 일별 변화 산포도



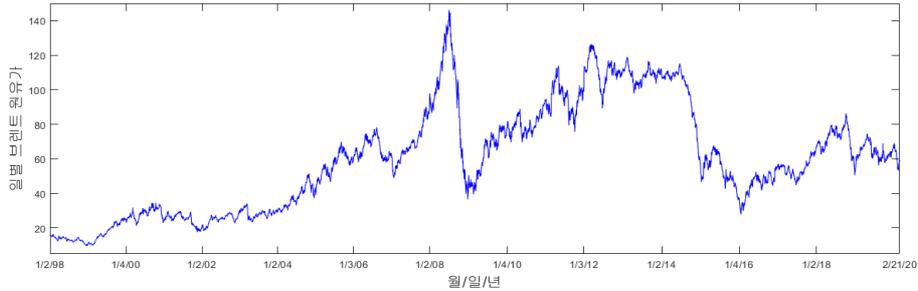
주: 1996년 1월 2일부터 2020년 2월 21일 까지 배럴당 미국 달러로 표시된 두바이유 가격에 대해 다음 날 계산된 일별 변화를 전일 원유가에 대해 산포도로 [그림 1. 나]에 나타냈다.

[그림 1. 다] 두바이유 가격의 일별 변화



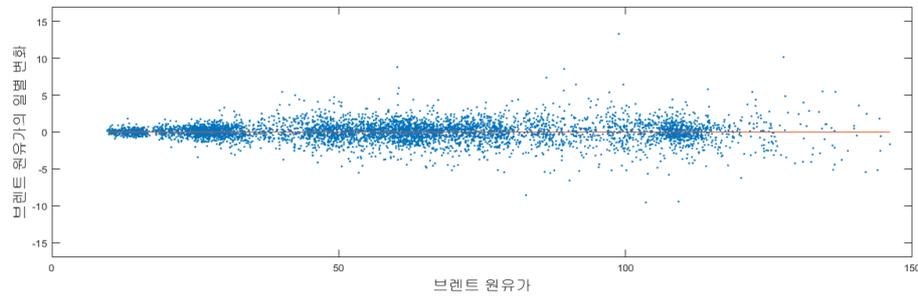
주: [그림 1. 다]는 1996년 1월 3일부터 2020년 2월 21일 까지 배럴당 미국 달러로 표시된 두바이유 가격의 일별 변화 시계열 자료를 보여주고 있다.

[그림 2. 가] 브렌트유의 일별 가격



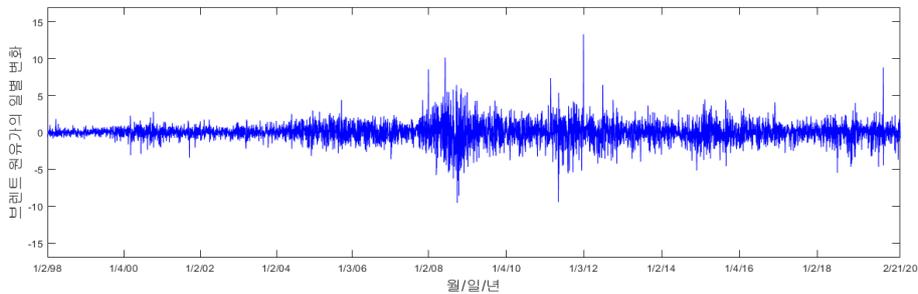
주: [그림 2. 가]에는 1998년 1월 2일부터 2020년 2월 21일 까지 배럴당 미국 달러로 표시된 브렌트유 가격의 일별 시계열 자료가 그려져 있다.

[그림 2. 나] 브렌트유 가격의 일별 변화 산포도



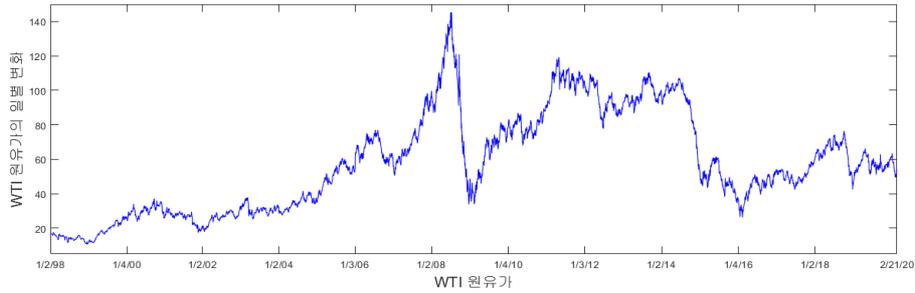
주: 1998년 1월 2일부터 2020년 2월 21일 까지 배럴당 미국 달러로 표시된 브렌트유 가격에 대해 다음 날 계산된 일별 변화를 전일 원유가에 대해 산포도로 [그림 2. 나]에 나타냈다.

[그림 2. 다] 브렌트유 가격의 일별 변화



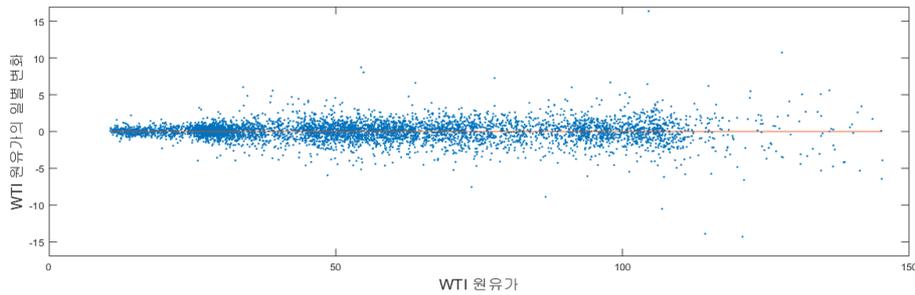
주: [그림 2. 다]는 1998년 1월 3일부터 2020년 2월 21일 까지 배럴당 미국 달러로 표시된 브렌트유 가격의 일별 변화 시계열 자료를 보여주고 있다.

[그림 3. 가] WTI유의 일별 가격



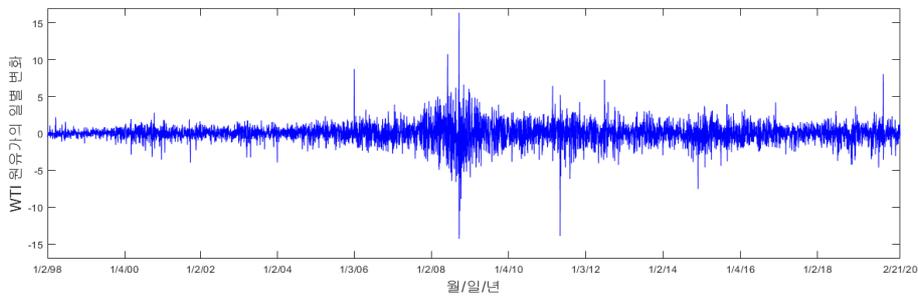
주: [그림 3. 가]에는 1998년 1월 2일부터 2020년 2월 21일 까지 배럴당 미국 달러로 표시된 WTI유 가격의 일별 시계열 자료가 그려져 있다.

[그림 3. 나] WTI유 가격의 일별 변화 산포도



주: 1998년 1월 2일부터 2020년 2월 21일 까지 배럴당 미국 달러로 표시된 WTI유 가격에 대해 다음 날 계산된 일별 변화를 전일 원유가에 대해 산포도로 [그림 3. 나]에 나타냈다.

[그림 3. 다] 브렌트유 가격의 일별 변화



주: [그림 3. 다]는 1998년 1월 3일부터 2020년 2월 21일 까지 배럴당 미국 달러로 표시된 WTI유 가격의 일별 변화 시계열 자료를 보여주고 있다.

양의 방향과 음의 방향 모두로 더 크게 변한 점들을 더 많이 볼 수는 있는데 이는 유가의 변동성이 유가의 수준에 따라 증가하는 것을 뜻할 수 있지만 유가가 장기적 평균값으로 회귀하는 경향이 있는지는 알기 어렵다. 그러므로 이러한 현상들이 통계적으로 의미가 있는지 여부를 알려줄 수 있는 적절한 모형을 추정하고 그 결과를 이용해 이러한 현상들의 존재 여부에 대해 좀 더 정확하게 논의하는 것이 필요하다. 이 논문에서 이용하는 확산과정 모형의 변동성 함수와 추세 함수의 추정 결과가 이들에 대한 답을 줄 수 있을 것이다.

세 번째 그림인 그림 [그림 1. 다], [그림 2. 다] 그리고 [그림 3. 다]는 각각 두바이유, 브렌트유 그리고 WTI유의 일별 변화의 시계열 자료를 그래프로 나타내고 있다. 이들 그림에서 공통적으로 유가의 변동 폭이 큰 기간과 작은 기간이 몰려 있는 변동성 군집 (volatility clustering) 현상이 관측되며 일별 변화가 상대적으로 큰 것들도 확인할 수 있다. 이는 변동성이 유가의 수준에 의존하기 때문일 수도 있고 경제 상황 등 국면에 따라 변동성 함수가 다르기 때문일 수도 있다.

### Ⅲ. 확산과정 모형과 추정 방법

이 연구에서는 세 종류의 대표적인 국제 원유가들의 움직임을 설명하는 모형으로 각 유가에 대해 다섯 개의 단일변수 시간균질 (time-homogeneous) 확산과정 (diffusion process)들을 추정한다. 이들 모형을 추정하기 위해 최우 추정 (maximum likelihood estimation)법을 이용한다.

#### 1. 확산과정 모형

1990년대 후반에 확산과정을 실제 데이터에 적용해 추정하는 여러 계량경제

〈표 2〉 원유가격의 확산과정 모형

	Model
Vasicek	$dX_t = (\alpha_0 + \alpha_1 X_t)dt + \sqrt{\beta_2} dW_t$
CIR	$dX_t = (\alpha_0 + \alpha_1 X_t)dt + \sqrt{\beta_2 X_t} dW_t$
CKLS	$dX_t = (\alpha_0 + \alpha_1 X_t)dt + \sqrt{\beta_2 X_t^{\beta_3}} dW_t$
GD-CEV	$dX_t = (\alpha_{-1} X_t^{-1} + \alpha_0 + \alpha_1 X_t + \alpha_2 X_t^2 + \alpha_3 X_t^3)dt + \sqrt{\beta_2 X_t^{\beta_3}} dW_t$
GD-GV	$dX_t = (\alpha_{-1} X_t^{-1} + \alpha_0 + \alpha_1 X_t + \alpha_2 X_t^2 + \alpha_3 X_t^3)dt + \sqrt{\beta_0 + \beta_1 X_t + \beta_2 X_t^{\beta_3}} dW_t$

주: 원유가의 모형을 위해 이용한 다섯 개의 다른 확산과정 모형이 제시되어 있다. GD-GV 모형은 나머지 네 개의 모형을 포함한다. 다른 모형들을 얻기 위한 모수의 제약도 함께 주어졌다.

학적 방법과 이론들이 개발되면서 확산과정을 이용한 실증 분석들이 많이 이루어졌는데 특히 단기이자율을 이용한 여러 연구 결과들이 있다. 이들 실증 분석에서는 이전에 이자율 기간구조 이론을 개발하기 위해 제안된 모형을 포함해 실증분석을 위해 새로 개발된 모형들을 이용했다. 예를 들면 Black and Scholes (1973), Merton (1973), Vasicek (1977), Dothan (1978), Cox, Ingersoll, and Ross (1985) (CIR), Chan, Karolyi, Longstaff, and Sanders (1992) (CKLS), 그리고 Ait-Sahalia (1996) 등이 있다.

이 논문에서는 단기 이자율의 움직임을 묘사하기 위해 Ait-Sahalia (1996)가 제안한 모형에서 추세 (drift) 함수에 3차 항을 추가해 Choi (2009)가 채택한 다음의 확산과정 모형을 주요 모형으로 해 원유가격의 움직임을 설명한다.

$$dX_t = (\alpha_{-1} X_t^{-1} + \alpha_0 + \alpha_1 X_t + \alpha_2 X_t^2 + \alpha_3 X_t^3)dt + \sqrt{\beta_0 + \beta_1 X_t + \beta_2 X_t^{\beta_3}} dW_t, \quad (1)$$

위 모형은 일반적인 형태의 추세 함수와 변동성 (volatility) 함수를 가지고 있어 지금까지 여러 금융 변수나 현물들의 가격을 모형화하기 위해 제안된 거의 모든 확산과정 모형들을 포함하고 있다. 따라서 이를 GD-GV (general

drift-general variance) 모형이라 부르고 GD-GV 모형에 포함되며 기존에 제안된 다른 모형들 중에서 대표적인 Vasicek, CIR, CKLS 모형들과, GD-GV에서 변동성 함수 부분이 CEV (constant elasticity of volatility) 형태인, GD-CEV 모형을 함께 추정한 후 비교할 것이다. <표 2>에는 본 연구에서 추정하는 모형들이 제시되어 있다.

원유와 같은 현물들의 가격이 이자율과 같이 가격이 너무 낮거나 너무 높으면 원유 시장의 수요나 공급이 조정되며 가격이 조정되어 장기적인 평균 수준으로 회귀한다고 주장하는 연구들이 있다 (Dixit and Pindyck (1994) 그리고 Smith and McCardle (1999)). GD-GV 모형은 일반적인 형태의 추세 함수를 가지고 있어 원유 가격이 장기적인 평균 수준으로 회귀하는지 여부뿐만 아니라 원유 가격의 수준에 따라 다른 속도로 회귀하는지 또한 확인할 수 있다. 모형 (1)의 변동성 함수는 CEV 형태에 상수항과 일차항을 추가한 것으로 원유 가격의 변동성이 가격이 늘어날 때 증가하는지 여부뿐만 아니라 가격이 낮을 때 또한 증가하는지 설명할 수 있다.

GD-GV 모형과 여기에 포함되는 모형들의 모수에 대한 제약이 필요하다. 아래에 주어진 제약식들은 Ait-Sahalia (1996)의 다섯 가지 조건들을 만족하도록 GD-GV 모형에 적절하게 구한 것이다.

$$\beta_0 \geq 0 (\beta_0 = 0 \text{ 그리고 } 0 < \beta_3 < 1 \text{ 이면 } \beta_2 > 0, \quad (2)$$

$$\text{또는 } \beta_0 = 0 \text{ 이고 } \beta_3 > 1 \text{ 이면 } \beta_1 > 0),$$

$$\beta_3 > 1 \text{ 이거나 } \beta_1 = 0 \text{ 이면 } \beta_2 > 0, \text{ 그리고} \quad (3)$$

$$0 < \beta_3 < 1 \text{ 이거나 } \beta_2 = 0 \text{ 이면 } \beta_1 > 0,$$

$$\alpha_3 \leq 0 (\text{그리고 } \alpha_3 = 0 \text{ 이면 } \alpha_2 \leq 0 \text{ 또는 } \alpha_3 = \alpha_2 = 0 \text{ 이면 } \alpha_1 < 0), \quad (4)$$

그리고

$$\alpha_{-1} > 0 \text{ 그리고 } 2\alpha_{-1} \geq \beta_0 \geq 0 \text{ 또는} \quad (5)$$

$$\alpha_{-1} = 0, \alpha_0 > 0, \beta_0 = 0, \beta_3 > 1 \text{ 그리고 } 2\alpha_0 \geq \beta_1 > 0$$

조건 (2)와 (3)은 각각 0과 무한 주변에서  $X_t$ 의 분산이 양의 값을 갖도록 하기 위한 것이다. 조건 (4)는  $X_t$ 가 무한의 값을 갖지 않도록 하는 것으로서  $X_t$ 가 큰 값을 가질 때 장기적인 평균 수준으로 회귀하도록 하는 역할을 한다. 원유 가격  $X_t$ 가 0의 값을 갖지 않도록 하는 것이 조건 (5)이다.

## 2. 최우추정법

위에서 소개된 확산과정 모형들을 추정하기 위해 가장 효율적(efficient)인 최우추정 (maximum likelihood estimation)법을 이용한다. 이를 위해서는 관측 자료의 우도를 알아야 한다. 확산과정 모형은 연속시간 모형이지만 우리가 관측할 수 있는 원유 가격 자료는 일별 자료와 같이 이산적이다. 즉, 관측치가  $n+1$ 개인 원유가  $X_t$ 는  $t = i\Delta, i = 0, 1, \dots, n$ 시점에만 관측이 된다. 그리고 우리는 일별 자료를 사용하므로 연속된 두 관측치 사이의 거리  $\Delta = \frac{1}{252}$ 이다<sup>2)</sup>. 그렇지만 확산과정의 마르코프 (Markov) 성질과 베이스 (Bayes) 정리를 이용해 관측된 자료의 우도함수를 구하고 로그를 취한 로그우도함수를 구하면 다음과 같이 나타낼 수 있다<sup>3)</sup>

$$l_n(\theta) = \sum_{i=1}^n \ln [p_X(\Delta, X_{i\Delta} | X_{(i-1)\Delta}; \theta)] \quad (6)$$

여기에서  $p_X(\Delta, X_{i\Delta} | X_{(i-1)\Delta}; \theta)$ 는 확산과정  $X_t$ 의 전이확률밀도 함수로  $X_t$ 가  $t = (i-1)\Delta$  때  $x_0$  값을 가지면 다음 시점인  $t = i\Delta$  때  $X_t$ 가  $x$  값을 가

2) 유가별로 연중 원유의 거래가 이루어지는 날들의 수가 조금씩 차이가 있지만 일별 자료를 사용하는 경우 거래가 이루어지지 않는 주말과 공휴일을 제외한 날수를 보통 252로 하는 관례를 따라 한 해의 거래일 수를 모두 252일로 통일했다.

3) 첫 관측치의 우도 값은 전체 우도값에 미치는 영향이 작기에 무시했다. 이에 대한 좀 더 자세한 설명은 Ait-Sahalia (2002)나 Choi (2020)을 참고하기 바란다.

질 조건부 확률밀도함수  $p_X(\Delta, x|x_0; \theta)$ 이다. 따라서  $X_t$ 의 전이확률밀도 함수만 알면 최우추정법을 이용해 확산과정 모형의 모수들을 추정할 수 있다. 그런데 문제는 Vasicek, CIR, Black-Scholes-Merton 모형들과 같이 몇 가지 예를 제외하고는 단일변수 모형이라고 하더라도 확산과정 모형의 전이확률밀도 함수는 알려져 있지 않다.

이러한 문제를 해결하기 위해 Ait-Sahalia (2002)는 임의의 단일변수 시간균질 확산과정의 전이확률밀도 함수를 매우 정확하게 근사시키는 방법을 개발했다. 그는 이 방법을 Ait-Sahalia (2008)에서 다변수 시간균질 확산과정의 로그 전이확률밀도 함수를 근사시키는 방법으로 발전시켰다. Ait-Sahalia (2002)에서 제안한 허마이트 (Hermite) 전개 방식으로 우리가 고려하는 확산과정들의 근사적 전이확률밀도 함수를 구할 수도 있으나 변동성 함수가 CEV나 GV와 같은 경우에는 더 복잡하거나 쉽지 않다. 그래서 본 논문에서는, 최우추정법을 이용하기 위해서는 로그 우도함수만 알면 되기 때문에 Ait-Sahalia (2008)가 확장시킨 방법을 이용해 확산과정들의 근사적 로그전이확률밀도 함수를 구했다<sup>4)</sup>. 이 방법에 대해 간단히 설명하면 다음과 같다.

먼저 근사적 로그전이확률밀도 함수를 다음 식과 같이 추론한다. 이러한 추론식은 막연한 추측이 아니며 허마이트 전개 방식으로 구한 전이확률밀도 함수에 로그를 취하면 얻어질 수 있는 형태를 생각해 구한 식이다.

$$l_X^{(K)}(\Delta, x|x_0; \theta) = -\frac{1}{2} \ln(2\pi\Delta) - D_v(x; \theta) + \frac{C_X^{(-1)}(\Delta, x|x_0; \theta)}{\Delta} \quad (7)$$

$$+ \sum_{k=0}^K C_X^{(k)}(\Delta, x|x_0; \theta) \frac{\Delta^k}{k!}$$

4) Ait-Sahalia (1999)는 Ait-Sahalia (2002)에서 제시한 허마이트 전개 방식이 아닌 다른 좀 더 편하게 근사적 전이확률밀도 함수를 구하는 방식을 이용해 여러 확산 과정들의 근사적 전이확률밀도 함수를 정확하게 구할 수 있다는 것을 보였다. Ait-Sahalia (2008)가 개발한 방법은 사실 허마이트 전개 방식으로 구한 것과 같다.

여기에서  $D_v(x;\theta) = \ln[\sigma(x;\theta)]$ 이며  $\Delta$ 의 K차 급수 전개인 이 식에서 각 계수들  $C_X^{(k)}(\Delta, x|x_0;\theta)$ ,  $k=-1,0,1,\dots,K$ 만을 구하면 된다. 이들을 구하기 위해 식 (7)을 확산과정의 로그 전이확률밀도 함수가 만족하는 콜모고로프 (Kolmogorov) 편미분 방정식에 대입한 후  $\Delta$ 에 대해 같은 차수끼리의 계수를 비교하면 아래와 같이 각 계수  $C_X^{(k)}(\Delta, x|x_0;\theta)$ 가 만족하는 미분 방정식을 얻을 수 있다.

$$\begin{aligned} 2C_X^{(-1)}(\Delta, x|x_0;\theta) &= -\sigma^2(x;\theta) \left[ \frac{dC_X^{(-1)}(\Delta, x|x_0;\theta)}{dx} \right]^2 \\ G_X^{(0)}(\Delta, x|x_0;\theta) &= -\sigma^2(x;\theta) \frac{dC_X^{(-1)}(\Delta, x|x_0;\theta)}{dx} \frac{dC_X^{(0)}(\Delta, x|x_0;\theta)}{dx} \\ G_X^{(k)}(\Delta, x|x_0;\theta) &= C_X^{(k)}(\Delta, x|x_0;\theta) \\ &\quad - \frac{1}{k} \sigma^2(x;\theta) \frac{dC_X^{(-1)}(\Delta, x|x_0;\theta)}{dx} \frac{dC_X^{(k)}(\Delta, x|x_0;\theta)}{dx}, \quad k \geq 1. \end{aligned}$$

위 미분 방정식 들에서  $G_X^{(k)}(\Delta, x|x_0;\theta)$ ,  $k=0,1,\dots$ 는 모두 확산과정의 추세 함수와 변동성 함수 그리고  $C_X^{(h)}(\Delta, x|x_0;\theta)$ ,  $h \leq k-1$ 의 식이다. 확산과정을 변환해 변동성 함수를 1로 만들 수 있다면 위 미분 방정식들을 풀어  $C_X^{(k)}(\Delta, x|x_0;\theta)$ 들을 구체적으로 구할 수 있다. 그런데 우리가 고려하는 모형들 중에서 GD-GV 모형은 이것이 가능하지 않기 때문에 Ait-Sahalia (2008)가 제안한 방식대로  $C_X^{(k)}(\Delta, x|x_0;\theta)$ 와  $G_X^{(k)}(\Delta, x|x_0;\theta)$ ,  $k=0,1,\dots$  모두를  $x$ 에 대해  $x_0$ 주변으로 테일러 (Taylor) 급수 전개를 해 각 미분 방정식에서  $(x-x_0)$ 의 같은 차수들끼리 비교해 테일러 급수 전개의 계수를 구해야 한다. 이렇게  $C_X^{(k)}(\Delta, x|x_0;\theta)$ 를  $j_k$ 차수까지 구한 식을  $C_X^{(j_k, k)}(\Delta, x|x_0;\theta)$ 라 하고 이를 식 (7)에 대입하면 다음과 같이 근사적 로그전이확률밀도 함수를 구할 수 있다5).

$$\begin{aligned} \tilde{l}_X^{(K)}(\Delta, x|x_0; \theta) = & -\frac{1}{2} \ln(2\pi\Delta) - \ln[\sigma(x; \theta)] + \frac{C_X^{(-1, j_{-1})}(\Delta, x|x_0; \theta)}{\Delta} \\ & + \sum_{k=0}^K C_X^{(k, j_k)}(\Delta, x|x_0; \theta) \frac{\Delta^k}{k!} \end{aligned}$$

그리고 이 식을 식 (6)에 대입하면 우리가 관측한 모든 자료에 대한 근사적 로그우도함수 값을 구할 수 있다.

$$l_n^{(K)}(\theta) \equiv \frac{1}{n} \sum_{i=1}^n \tilde{l}_X^{(K)}(\Delta, X_{i\Delta} | X_{(i-1)\Delta}; \theta)$$

이 식을 극대화 하는 모수가 최우추정 값이다. Vasicek과 CIR 모형은 참 전이확률밀도 함수가 알려져 있지만 일관성을 위해 다른 모형들과 마찬가지로 근사적 로그전이확률밀도 함수를 이용해 최우추정량을 구했다<sup>6)</sup>. 그리고 2장에서 소개한 유가 자료는 1 배럴당 미국 달러로 표시된 일별 유가 자료인데, 확산과정 모형들을 추정할 때는 이 수치들을 100으로 나눠 유가를 0.01 배럴당 미국달러로 나타낸 후에 이들을 이용해 모형을 추정했다<sup>7)</sup>.

- 
- 5) 좀 더 자세한 설명은 Ait-Sahalia (2008)이나 Choi (2020)을 참고하기 바란다.
- 6) Ait-Sahalia (2008)에서 이론적으로 그리고 다양한 확산과정들에 대해 몬테카를로 시뮬레이션으로 보였듯이 이렇게 구한 근사적 로그전이확률밀도 함수는 K=1 또는 2인 경우에도 참 로그전이확률밀도 함수를 매우 정확하게 근사시킬 수 있으며 근사적 로그전이확률밀도 함수를 이용해 구한 최우추정 값은 참 값에 매우 가깝다.
- 7) 이는 급수 전개식으로 표현된 근사적 로그전이확률밀도 함수를 이용해 추정할 때 좀 더 정확한 근사식을 얻기 위한 것이며 이를 고려해 추정 결과를 해석하면 된다. 회귀 분석 모형에서 변수의 측정 단위를 바꾸면 추정 값은 달라지지만 추정결과의 해석에는 영향을 주지 않는 것과 같다. 또한 확산과정 모형 추정에 많이 이용된 단기 이자율 자료들은 보통 %로 표시되어 제공되는데 단기 이자율을 이용해 확산과정을 추정할 때도 이 값들을 100으로 나누어 추정한다.

#### IV. 추정 결과와 논의

다섯 개의 확산과정들의 모수들을 두바이유, 브렌트유 그리고 WTI유의 일별 자료를 이용해 추정한 결과가 각각 <표 3>, <표 4> 그리고 <표 5>에 정리되어 있다. 그리고 각 모형의 추정 결과들을 이용해 추세 함수와 변동성 함수를 이들의 95% 신뢰밴드와 함께 그려 각각 [그림 4], [그림 5] 그리고 [그림 6]에 제시했다. 2장에서 세 가지 다른 유가들의 기초 통계량과 그래프들을 보며 논의 했듯 이들이 비슷한 성질을 가지고 있고 전반적으로 시간에 따라 유사하게 움직이는 것을 확인할 수 있었는데 확산과정 모형들을 추정한 결과 또한 비슷하고 일관된 면이 많이 있었다.

<표 3>의 두바이유 추정 결과를 먼저 보면 마지막 두 줄에 계산된 정보기준 AIC와 BIC<sup>8)</sup>를 이용해 모형들을 비교했을 때 CKLS 모형은 이 모형이 포함하는 Vasicek 모형과 CIR 모형은 물론 이보다 더 일반적인 추세함수나 변동성 함수를 가진 GD-CEV 모형과 GD-GV 모형보다 더 낫다는 결과를 얻었다. 이는 GD-CEV 모형과 GD-GV 모형의 추세 함수의 모수들이 어떤 것도 통계적으로 유의하지 않다는 사실 또한 설명한다. 그리고 GD-GV 모형에서 변동성 함수에 새로 추가한 1차식의 모수  $\beta_0$ 와  $\beta_1$ 이 매우 작은 값을 갖고 통계적으로도 유의하지 않다. 또한 가장 일반적인 모형인 GD-GV 모형을 여기에 포함되는 다른 모형들 각각과 비교하기 위해 우도비검정 통계량을 계산한

8)  $AIC = -2\ln L(x, \hat{\theta}_n) + 2d$  그리고  $BIC = -2\ln L(x, \hat{\theta}_n) + \ln(n)d$ 이며 여기에서  $\ln L(x, \hat{\theta}_n)$ 는 최대 로그우도 값이고 d는 각 모형에서 추정하는 모수들의 수 그리고 n은 전체 관측치의 수에서 1을 뺀 값이다. 정보기준은 서로 경쟁하는 모형들을 비교하기 위한 것으로, 각 모형이 주어진 자료를 완벽히 설명하지 못함으로써 생기는 정보의 손실을 비교해 여러 모형들 중에서 그러한 정보 손실이 상대적으로 적은 모형을 선택할 수 있도록 해주는 기준이다.

결과가 밑에서 세 번째 줄에 있는데, 여기에서 볼 수 있듯 CKLS 모형이나 GD-CEV 모형이 기각되지 못했다. 이는 CKLS 모형보다 더 일반적인 형태의 추세 함수나 변동성 함수가 두바이유의 움직임을 설명하는데 도움이 되지 않는다는 것을 뜻한다. 그렇지만 Vasicek 모형과 CIR 모형 보다는 좀 더 유연한  $X_t$ 의 변동성 함수를 갖는 CKLS 모형이 두바이유의 움직임을 설명하기에 더 적절하다는 것을 나타낸다. 브렌트유와 WTI유에서 얻은 모수들의 추정 값에는 약간 차이가 있기는 했지만, 이들 원유가에 대해서도 비슷한 결과를 얻을 수 있었다. 흥미롭게도 이러한 결과는 이자율 (Choi (2015b))이나 환율 (Choi and Lee (2020))에 대한 추정 결과에서는 GD-GV 모형이 가장 우월했다는 사실과는 다르다. 즉 자료의 동태적인 성질을 설명할 때 더 일반적인 모형을 이용하는 것이 항상 좋은 것은 아니라는 사실을 보여주고 있다.

GD-GV 모형의 변동성 함수에 추가된 모두 모수  $\beta_0$ 와  $\beta_1$ 를 제외하고는 모든 모형들에서 변동성 함수의 모수들이 통계적으로 매우 유의했다. 반면에 CKLS 모형의  $\alpha_0$ 를 제외하고는 추세 함수의 나머지 모든 모수들은 5% 유의수준에서 통계적으로 유의하지 않은 결과를 얻었다. 이는 두바이유의 경우 추세함수 보다 변동성 함수가 유가의 움직임을 설명하는데 더 중요하다는 것을 뜻한다.

CKLS 모형이 가장 선호되므로 이모형의 추정 결과를 중심으로 추정된 모수들의 값이 의미하는 것을 살펴보자. 추세함수의 모수  $\alpha_0$ 은 통계적으로 유의하지만  $\alpha_1$ 은 통계적으로 유의하지 않다. 그렇지만 이 값들을 이용해 두바이유 가격의 장기적인 평균수준을 계산하면  $-\alpha_0/\alpha_1 = 1.0513$  즉 배럴당 105.13달러로 관측된 두바이유의 최댓값 140.70과 최솟값 9.45 사이 값이기는 하지만 표본평균값인 54.84과는 큰 차이를 보인다. 수렴하는 속도를 나타내는 모수  $\alpha_1$ 의 추정 값을 이용해 장기적인 평균수준으로 회귀하기까지 걸리는 시간의 절반인 반감기 (half-life)<sup>9)</sup>를 계산하면 17.77년으로 다소 긴 것을 알 수

9) 확산과정의 추세함수가  $\alpha_0 + \alpha_1 X_t$ 와 같이 선형함수인 경우  $\alpha_1$ 이 음수면  $X_t$ 는 장기적인 평균수준  $-\alpha_0/\alpha_1$ 으로 회귀하며 장기적인 평균수준으로 회귀하기까지 걸리는 시간의 절반인 반감기는  $-\ln(2)/\alpha_1$ 인데 일별 자료를 사용하므로 반감기는  $-252\ln(2)/\alpha_1$ 다.

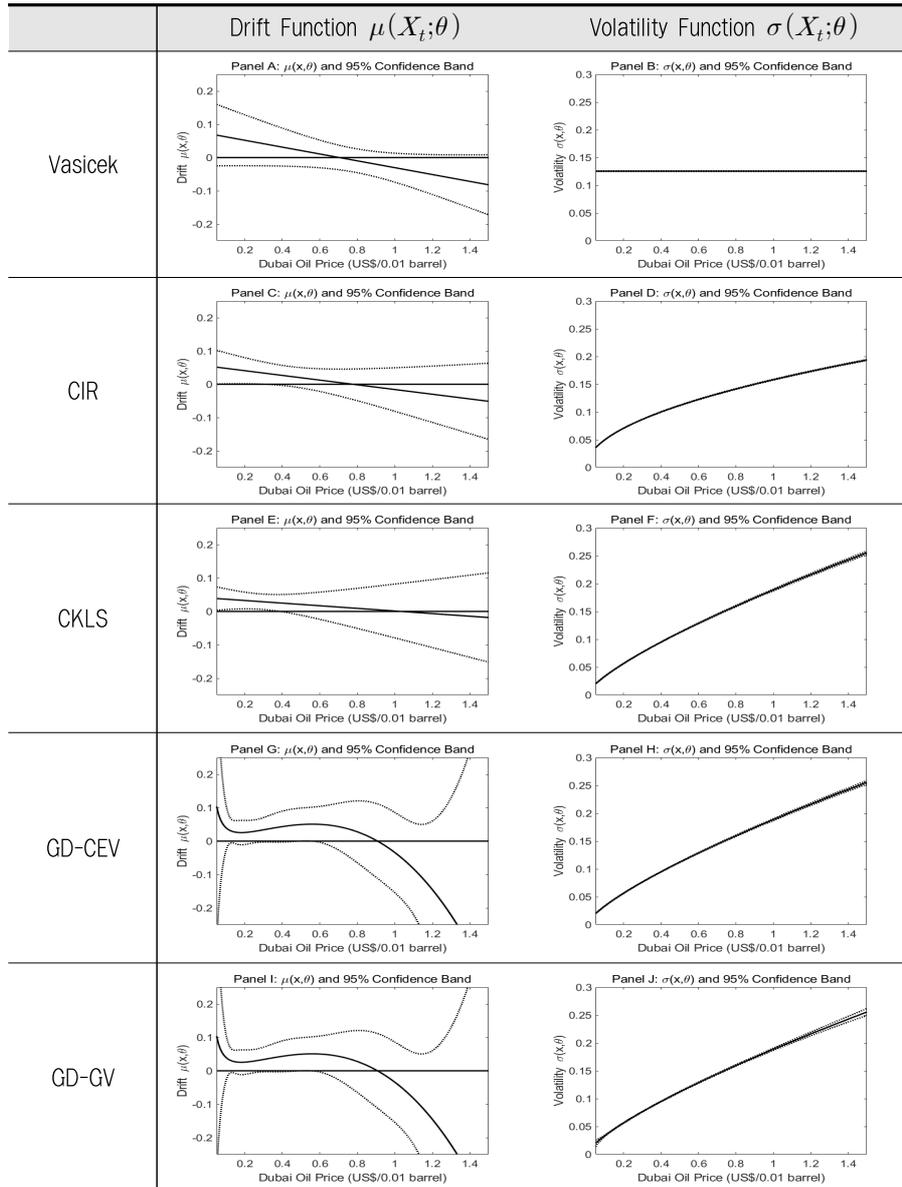
있다. 그런데 이러한 결과는 통계적으로 유의한  $\alpha_1$  값을 얻을 수 없었기 때문에 나오는 결과라 생각된다.

〈표 3〉 두바이유 추정 결과

$\theta$	Vasicek	CIR	CKLS	GD-CEV	GD-GV
$\alpha_{-1}$ (std error)	0 -	0 -	0 -	0.0077 (0.024)	0.0077 (0.024)
$\alpha_0$ (std error)	0.073 (0.050)	0.055 (0.028)	0.041** (0.020)	-0.064 (0.33)	-0.064 (0.33)
$\alpha_1$ (std error)	-0.10 (0.059)	-0.071 (0.053)	-0.039 (0.056)	0.28 (1.40)	0.28 (1.40)
$\alpha_2$ (std error)	0 -	0 -	0 -	-0.078 (2.14)	-0.078 (2.14)
$\alpha_3$ (std error)	0 -	0 -	0 -	-0.18 (1.02)	-0.18 (1.02)
$\beta_0$ (std error)	0 -	0 -	0 -	0 -	0.000000 (0.00033)
$\beta_1$ (std error)	0 -	0 -	0 -	0 -	0.0000021 (0.010)
$\beta_2$ (std error)	0.016** (0.000067)	0.025** (0.00013)	0.036** (0.00033)	0.036** (0.00033)	0.036** (0.0099)
$\beta_3$ (std error)	0 -	1 -	1.50** (0.0090)	1.50** (0.0091)	1.50** (0.16)
log-lik	30496.69	32566.54	32858.88	32859.56	32859.56
LR test (p-value)	GD-GV vs Vasicek 0.000	GD-GV vs CIR 0.000	GD-GV vs CKLS 0.93	GD-GV vs GD-CEV 1	-
AIC	-60987.38	-65127.08	-65709.76	-65705.12	-65701.12
BIC	-60967.23	-65106.93	-65682.89	-65658.11	-65640.67

주: 모수들의 최우추정 값과 이들의 표준오차가 괄호 안에 있다. 유의수준 5%에서 통계적으로 유의한 모수의 추정 값을 얻은 경우에는 추정 값 옆에 \*\*표시를 했다. 모수들의 추정 결과 다음에는 각 모형에 대한 최대 로그우도함수의 값이 있고, GD-GV 모형을 다른 모형들의 가정에 대한 우도비 검정 결과의 p-값이 있다. 마지막 두 열에는 각각 정보기준 AIC와 BIC가 제시되어 있다.

[그림 4] 두바이유의 추세와 변동성 함수



주: 최우추정 값을 이용해 각 모형에 대해 추세 함수와 변동성 함수 그리고 이들의 95% 신뢰밴드를 관측된 유가의 범위에서 그렸다.

다른 한편 추세 함수의 모수들이 모두 통계적으로 유의한 값을 얻지는 못했지만, 관측된 유가의 모든 범위에서 유가가 정말로 장기적인 평균수준으로 회귀하는 성질이 없는지 여부를 알기위해 추세 함수와 이의 95% 신뢰밴드를 그려 확인해보는 것이 필요하다. 두바이유를 다섯 가지 확산과정들에 적용해 추정한 결과를 이용해 각각에 대해 추세 함수와 95% 신뢰밴드를 [그림 4]의 두 번째 열에 그렸다. CKLS의 추세 함수 그림에서 알 수 있듯 유가가 배럴당 20달러 근처의 값들에서는 95% 신뢰밴드가 x축을 포함하지 않고 양의 값을 갖기에 유가가 더 높은 수준으로 늘어나는 경향이 있음을 확인할 수 있다. 이는 Vasicek 모형을 제외한 나머지 모형들에서도 유가가 낮을 때 회귀성이 있는 것으로 나타났다. GD-CEV 모형과 GD-GV 모형의 추세함수 모수들은 어떤 것도 통계적으로 유의하지 않았지만 추정 결과들을 이용해 그린 추세함수의 모습에서는 통계적으로 0과 다른 구간들이 있었다는 사실에 유의하자. 그렇지만 모든 모형에서 유가가 배럴당 50달러에서 60달러 범위를 넘어서면 95% 신뢰밴드가 x축을 포함해 회귀성이 없음을 알 수 있다. 이와 같이 유가가 높을 때는 장기적인 평균 수준으로 회귀한다는 증거를 찾지 못한 이유는 원유에 대한 수요는 전 세계적이기 때문에 세계의 경제 상황에 따라 결정되지만 공급은 주요 산유국들에 의해 결정되기 때문일 수 있다. 경제 상황이 나빠 원유에 대한 수요가 낮아져 가격이 내려가면 소수의 주요 산유국들이 수익성을 높이기 위해 공급을 줄임으로써 가격을 올리려는 유인이 있을 것이다. 그렇지만 경제 상황이 좋아져 유가가 올라가면 원유 공급자들은 수익을 많이 내고 있는 상황에서 일부러 공급을 더 늘려 가격을 낮출 유인은 적기 때문에 유가가 높을 때는 하방경직성이 있다는 결과를 얻은 것이 아닌가 생각된다.

위와 같은 결과는 브렌트유와 WTI유에서도 비슷하게 얻어진다. 브렌트유와 WTI유에서 CKLS 모형의 결과를 보면 각각 장기적 평균 수준은 배럴당 94.12달러와 78달러로 여전히 각각의 표본 평균 값들 보다는 컸다. 반감기는 10.19와 6.93년으로 두 경우 모두 두바이유보다는 작게 나왔다. 그렇지만 이 두 경우에도  $\alpha_0$ 는 통계적으로 유의하지만  $\alpha_1$ 은 그렇지 않았다. 이들 유가를

이용해 구한 추정 값들로 [그림 5]와 [그림 6]의 두 번째 열에 그린 추세함수를 보면 두바이유와 마찬가지로 Vasicek 모형을 제외하고는 낮은 유가에서 모두 장기적인 평균 수준으로 회귀하는 증거를 찾을 수 있었으며 이 구간을 지나면 모든 경우에 95% 신뢰밴드가 x축을 포함해 추세함수가 0과 다르지 않다는 것을 알 수 있다.

유가와 같은 현물들은 시장의 수요와 공급에 따라 장기적인 평균 수준으로 회귀하는 성질이 있다고 주장된다. 그리고 Dixit and Pindyck (1994)에서 설명하듯 원유 생산 비용과 원유 가격의 장기적인 평균수준 그리고 장기적인 평균 수준으로 회귀하는 속도는 원유 개발 사업의 수익성을 결정하는데 중요한 역할을 한다. 그러므로 원유가가 장기적인 평균 수준으로 돌아오는 성질이 있는지 여부와 그렇다고 하면 이들 모수를 정확하게 추정하는 것이 중요하다. 그런데 적어도 본 연구에서 살펴본 기간의 주요 원유가들에서는 유가가 낮을 때는 장기적 평균수준으로 회귀하는 경향을 확인할 수 있었으나 이후의 유가에서는 그러지 못했다. 또한 추세함수의 모수들이 모두 정확하게 추정된 것이라 할 수 없어 이들로부터 구한 장기적인 평균 수준도 정확한 값이라고 말할 수는 없다.

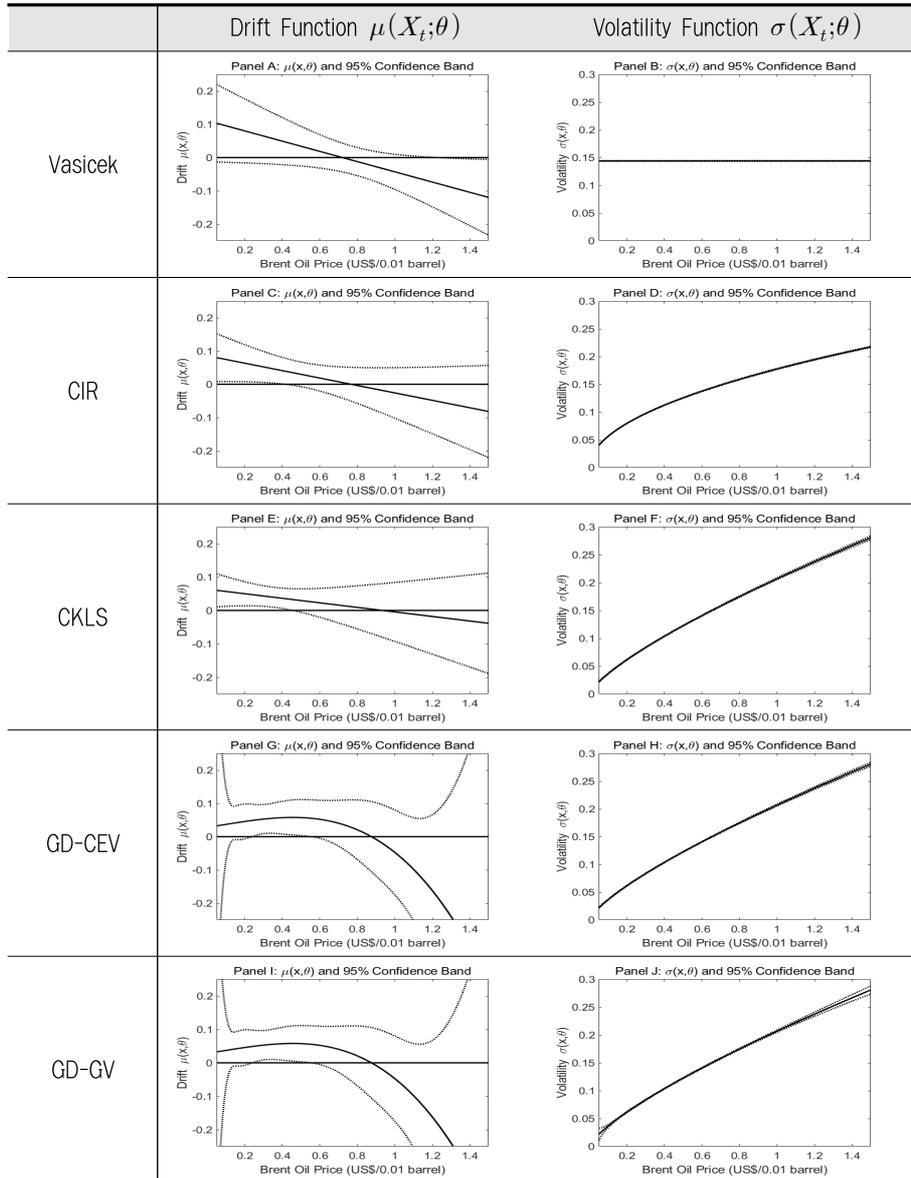
다음으로 변동성함수의 추정 결과를 살펴보자. 두바이유의 CKLS 모형에서 CEV 형태의 변동성함수에 있는 모수들  $\beta_2$ 와  $\beta_3$ 모두가 통계적으로 매우 유의하다. 확산과정 모형의 변동성 함수가  $\sigma(X_t; \theta)$ 일 때 변동성의  $X_t$ 에 대한 탄력성은  $[d\sigma(X_t; \theta)/\sigma(X_t; \theta)]/(dX_t/X_t)$ 로 계산하는데 CKLS 모형의 변동성 함수는  $\sqrt{\beta_2 X_t^{\beta_3}}$ 로 유가  $X_t$ 에 대한 변동성의 탄력성은  $\beta_3/2$ 이다. 따라서  $\beta_3$ 의 추정 값은 1.5로 유가가 1% 증가하면 변동성은 이보다 작은 비율인 0.75% 늘어난다는 것을 뜻한다. GD-CEV와 GD-GV 모형에서  $\beta_0$ 와  $\beta_1$ 의 추정 값들은 사실상 0의 값을 갖고 통계적으로도 유의하지 않지만 나머지 모든 모형들의 변동성 함수와 관련된 모수들은 모두 p-값이 매우 작아 통계적으로 아주 유의하다. 추세함수의 모수들이 통계적으로 유의하지 않은 편이었던

〈표 4〉 브렌트유 추정 결과

$\theta$	Vasicek	CIR	CKLS	GD-CEV	GD-GV
$\alpha_{-1}$ (std error)	0 -	0 -	0 -	0.0000000 (0.032)	0.000023 (0.032)
$\alpha_0$ (std error)	0.11 (0.063)	0.086** (0.040)	0.064** (0.028)	0.028 (0.42)	0.028 (0.42)
$\alpha_1$ (std error)	-0.15** (0.075)	-0.11 (0.069)	-0.068 (0.066)	0.088 (1.63)	0.088 (1.63)
$\alpha_2$ (std error)	0 -	0 -	0 -	0.048 (2.35)	0.048 (2.35)
$\alpha_3$ (std error)	0 -	0 -	0 -	-0.21 (1.081)	-0.21 (1.084)
$\beta_0$ (std error)	0 -	0 -	0 -	0 -	0.0000000 (0.00059)
$\beta_1$ (std error)	0 -	0 -	0 -	0 -	0.00000020 (0.016)
$\beta_2$ (std error)	0.021** (0.000089)	0.032** (0.00017)	0.043** (0.00037)	0.043** (0.00038)	0.043** (0.015)
$\beta_3$ (std error)	0 -	1 -	1.50** (0.011)	1.51** (0.011)	1.51** (0.20)
log-lik	26688.21	28128.14	28338.28	28338.95	28338.95
LR test (p-value)	GD-GV vs Vasicek 0.000	GD-GV vs CIR 0.000	GD-GV vs CKLS 0.93	GD-GV vs GD-CEV 1	-
AIC	-53370.42	-56250.28	-56668.56	-56663.90	-56659.90
BIC	-53350.50	-56230.36	-56642.00	-56617.42	-56600.14

주: 모수들의 최우추정 값과 이들의 표준오차가 괄호 안에 있다. 유의수준 5%에서 통계적으로 유의한 모수의 추정 값을 얻은 경우에는 추정 값 옆에 \*\*표시를 했다. 모수들의 추정 결과 다음에는 각 모형에 대한 최대 로그우도함수의 값이 있고, GD-GV 모형을 다른 모형들의 가정에 대한 우도비 검정 결과의 p-값이 있다. 마지막 두 열에는 각각 정보기준 AIC와 BIC가 제시되어 있다.

[그림 5] 브렌트유의 추세와 변동성 함수



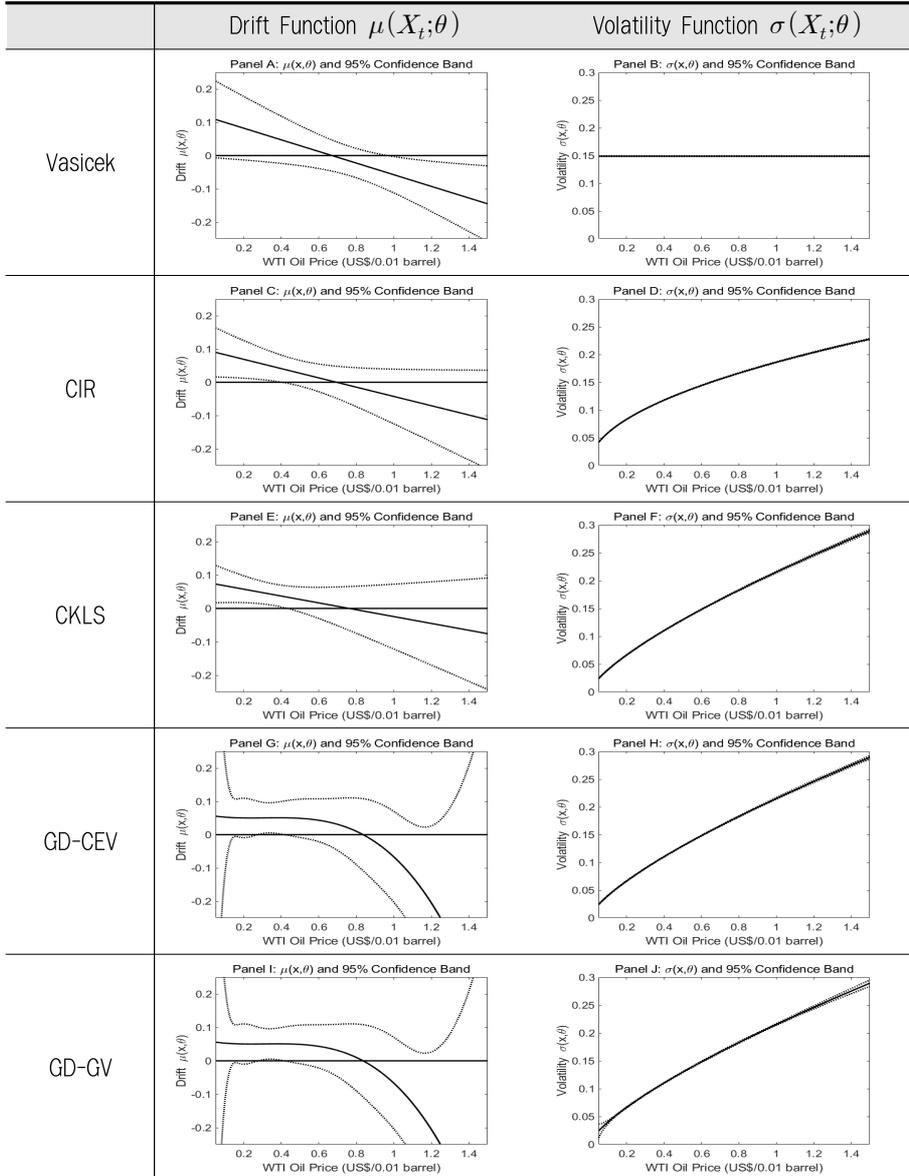
주: 최우추정 값을 이용해 각 모형에 대해 추세 함수와 변동성 함수 그리고 이들의 95% 신뢰밴드를 관측된 유가의 범위에서 그렸다.

〈표 5〉 WTI유 추정 결과

$\theta$	Vasicek	CIR	CKLS	GD-CEV	GD-GV
$\alpha_{-1}$ (std error)	0 -	0 -	0 -	0.000000 (0.043)	0.000013 (0.044)
$\alpha_0$ (std error)	0.12 (0.062)	0.097** (0.041)	0.078** (0.032)	0.060 (0.52)	0.060 (0.53)
$\alpha_1$ (std error)	-0.17** (0.075)	-0.14 (0.073)	-0.10 (0.074)	-0.10 (1.95)	-0.10 (1.97)
$\alpha_2$ (std error)	0 -	0 -	0 -	0.35 (2.71)	0.35 (2.73)
$\alpha_3$ (std error)	0 -	0 -	0 -	-0.37 (1.22)	-0.37 (1.23)
$\beta_0$ (std error)	0 -	0 -	0 -	0 -	0.000000 (0.00072)
$\beta_1$ (std error)	0 -	0 -	0 -	0 -	0.000014 (0.021)
$\beta_2$ (std error)	0.022** (0.000075)	0.035** (0.00016)	0.046** (0.00034)	0.046** (0.00034)	0.046** (0.020)
$\beta_3$ (std error)	0 -	1 -	1.46** (0.0091)	1.46** (0.0092)	1.46** (0.21)
log-lik	25848.43	27238.48	27405.87	27406.32	27406.32
LR test (p-value)	GD-GV vs Vasicek 0.000	GD-GV vs CIR 0.000	GD-GV vs CKLS 0.90	GD-GV vs GD-CEV 1	-
AIC	-51690.86	-54470.96	-54803.74	-54798.64	-54794.64
BIC	-51670.99	-54451.09	-54777.25	-54752.29	-54735.04

주: 모수들의 최우추정 값과 이들의 표준오차가 괄호 안에 있다. 유의수준 5%에서 통계적으로 유의한 모수의 추정 값을 얻은 경우에는 추정 값 옆에 \*\*표시를 했다. 모수들의 추정 결과 다음에는 각 모형에 대한 최대 로그우도함수의 값이 있고, GD-GV 모형을 다른 모형들의 가정에 대한 우도비 검정 결과의 p-값이 있다. 마지막 두 열에는 각각 정보기준 AIC와 BIC가 제시되어 있다.

[그림 6] WTI유의 추세와 변동성 함수



주: 최우추정 값을 이용해 각 모형에 대해 추세 함수와 변동성 함수 그리고 이들의 95% 신뢰밴드를 관측된 유가의 범위에서 그렸다.

것과 비교하면 이는 유가의 움직임을 설명하는데 추세함수 보다는 변동성 함수가 더 중요한 역할을 한다는 것을 말해준다.

추세함수의 경우와 마찬가지로 두바이유의 경우 추정된 모수들을 이용해 그린 모든 모형들의 변동성 함수와 이의 95% 신뢰밴드가 [그림 4]의 마지막 열에 제시되어 있다. 이 그림을 보면 추정결과를 보고 예상할 수 있듯, 유가의 변동성이 유가의 증가함수임을 알 수 있다<sup>10)</sup>. 그리고 좀 더 유연한 형태의 변동성 함수를 이용해 유가가 낮을 때 변동성이 증가하는지 여부를 알아보려고 했으나 이와 관련한  $\beta_0$ 와  $\beta_1$ 의 추정 값들이 모두 실질적으로 0이었기 때문에 그러한 증거를 찾을 수 없었다.

브렌트유와 WTI유의 변동성 함수 추정 결과도 두바이유와 별로 다르지 않았다. CKLS 모형을 이용해 변동성의 유가에 대한 탄력성을 보면 브렌트유는 0.75로 두바이유와 같고 WTI유는 0.73으로 두바이유보다 약간 작았지만 큰 차이는 없었다. 이들의 경우에도 모두  $\beta_0$ 와  $\beta_1$ 의 추정 값들이 실질적으로 0이었다. 추정 값들을 이용해 [그림 5]와 [그림 6]의 마지막 열에 그린 변동성 함수와 이의 95% 신뢰밴드를 보면 두바이유에서와 매우 비슷한 것을 확인할 수 있다. 따라서 브렌트유와 WTI유도 추세함수 보다는 변동성 함수가 이들의 움직임을 설명하는데 더 중요하다고 할 수 있다.

## V. 결론

이 연구의 목적은 유가의 움직임을 파악할 때 대표적으로 이용되는 두바이유, 브렌트유 그리고 WTI유의 일별 자료를 이용해 유가의 동태적 움직임을 잘 설명할 수 있는 확산과정을 찾는 것이다. 이를 위해 선행연구들에서 제안된 대부분의 추세 함수와 변동성 함수를 포함하는 일반적인 추세 함수와 변

10) 물론 Vasicek 모형의 변동성은 유가에 의존하지 않는 상수이기 때문에 상수로 나타내진다.

동성 함수를 갖는 확산과정인 GD-GV 모형 그리고 이에 포함되며 기존의 연구들에서 많이 활용된 Vasicek, CIR, CKLS 그리고 GD-CEV 확산과정 모형들을 유가 자료들에 적용해 추정했다. 최우추정법을 이용해 확산과정들을 추정했는데 이를 위해 필요한 로그 전이확률밀도 함수가 일반적으로 알려져 있지 않기 때문에 Ait-Sahalia (2008)가 개발한 방법을 이용해 근사적으로 구했다. 추정결과를 보면 모든 유가들에 대해 CKLS 모형이 유가의 움직임을 가장 잘 설명한다는 결과를 얻었다. CKLS 모형이 이를 포함하는 Vasicek 모형과 CIR 모형은 물론 이 모형을 포함하는 GD-GV 모형과 GD-CEV 모형보다 더 선호된다는 강한 증거들을 찾았다. 먼저 정보기준 AIC와 BIC 모두에 대해 CKLS가 더 선호된다는 것을 알 수 있었고 GD-GV 모형과 GD-CEV 모형에 추가된 모수들이 통계적으로 유의하지 않았다. 추세 함수 보다는 변동성 함수와 관련된 모수들이 통계적으로 매우 유의했는데 이는 유가의 움직임을 설명하는데 추세 함수 보다는 변동성 함수가 더 중요하다는 것을 뜻한다. 낮은 유가에서는 장기 평균값으로 회귀한다는 증거를 찾을 수 있었으나 중간 이상의 유가에서는 그러한 증거를 찾지 못했다. 변동성의 유가에 대한 탄력성은 약 0.75로 추정됐다. 본 논문의 의의는 아주 일반적인 형태의 확산과정을 대표적인 유가들에 적용하고 확산과정 추정에 적합한 방법론을 이용해 추정해 유가들의 동태적 변화를 가장 잘 설명하는 모형을 구했다는 데 있다. 일부의 이론적 또는 실증적 선행 연구들에서는 단일변수 확산과정뿐만 아니라 점프 확산과정이나 국면전환을 함께 고려한 확산과정 또는 2요인 이상을 포함한 확산과정 모형들도 함께 고려하고 있다. 이러한 요소들을 추가한 확산과정 모형들에 유가를 적용해 추정하는 것도 앞으로 이어질 연구들에서 진행할 예정이다. 또한 추정 결과를 이용해 유가 자료의 움직임에 적합한 모형을 찾은 후에 이 결과가 유가 선물이나 옵션 그리고 파생상품들의 가격을 계산하는데 주는 의미가 무엇인지 고찰해 보는 것 또한 흥미로운 연구 분야라 생각된다.

접수일(2021년 2월 9일), 수정일(2021년 3월 17일), 게재확정일(2021년 3월 26일)

◎ 참 고 문 헌 ◎

- 윤원철 2005. “확률과정 모형을 활용한 석유가격 추정 및 시사점,” 『에너지경제연구』, 제4권 제1호, pp. 39-61.
- 윤원철, 박호정 2005. “원유가격의 동태성 추정과 옵션가치 산정,” 『자원·환경경제연구』, 제14권 제4호, pp. 943-964.
- 정세록 2016. “국제원유시장가격의 확률과정 모형화에 대한 메타연구 및 실증연구,” 서울대학교 석사논문
- Ait-Sahalia, Y. 1996. “Testing Continuous-time Models of the Spot Interest Rate,” *Review of Financial Studies*, Vol. 9(2), pp. 385-426.
- \_\_\_\_\_ 2002. “Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed-form Approximation Approach,” *Econometrica*, Vol. 70(1), pp. 223-262.
- \_\_\_\_\_ 2008. “Closed-form Likelihood Expansions for Multivariate Diffusions,” *Annals of Statistics*, Vol. 36(2), pp. 906-937.
- Akaike, H. 1973. “Information Theory and an Extension of the Maximum Likelihood Principle, in B. N. Petrov and F. Csaki, eds. Budapest: Akademiai Kiado,” *Second International Symposium*, pp. 267-281.
- Askari, H. and N. Krichene 2008. "Oil price dynamics(2002-2006)", *Energy Economics*, Vol. 30, pp. 2134-2153.
- Bates, D.S. 1996. "Jumps and stochastic volatility: Exchange rate processes implicit in deutsche mark options," *Review of Financial Studies*, pp 69 - 107.
- Bernabe, A., Martina, E., Alvarez-Ramirez, J. and C. Ibarra-Valdez 2004. “A multi-model approach for describing crude oil price dynamics”, *Physica A*, Vol. 338, pp. 567-584.
- Black, F., and M. Scholes 1973. “The Pricing of Options and Corporate Liabilities,” *Journal of Political Economy*, Vol. 81(3), pp. 637-654.
- Brennan. M. J. and E. S. Schwartz 1985. “Evaluating national resource investments”, *The*

- Journal of Business, Vol. 58, No. 2, pp. 135-157.
- Chan, K., Karolyi, F. Longstaff, and A. Sanders 1992. "An Empirical Comparison of Alternative Models of the Short-Term Interest Rate," *Journal of Finance*, Vol. 47(3), pp. 1209-1227.
- Choi, S. 2009. "Regime-Switching Univariate Diffusion Models of the Short-Term Interest Rate," *Studies in Nonlinear Dynamics & Econometrics*, Vol. 13(1), Article 4.
- \_\_\_\_\_ 2013. "Closed-Form Likelihood Expansions for Multivariate Time-inhomogeneous Diffusions," *Journal of Econometrics*, Vol. 174(2), pp. 45-65.
- \_\_\_\_\_ 2015a. "Explicit Form of Approximate Transition Probability Density Functions of Diffusion Processes," *Journal of Econometrics*, Vol. 187, pp. 57-73.
- \_\_\_\_\_ 2015b. "Maximum Likelihood Estimation of Continuous-time Diffusion Models for Korean Short-Term Interest Rates", *Economic Analysis*, Vol. 21 (4), pp. 28-58.
- \_\_\_\_\_ 2019. "Approximate Transition Probability Density Function of a Multivariate Time-inhomogeneous Jump Diffusion Process in a Closed-Form Expression," Working paper, University of Seoul.
- \_\_\_\_\_ 2020. "Recent Development of Closed-form Approximate (Log-)Transition Probability Density Functions of Diffusion Processes," *Journal of Economic Theory and Econometrics*, 31 (1), 97-152.
- Choi, S. and Lee J. 2020. "Maximum Likelihood Estimation of Continuous-time Diffusion Models for Exchange Rates," *East Asian Economic Review*, 24(1), 61-87
- Cortazar, G. and E. S. Schwartz 1997. Implementing a real option model for valuing an undeveloped oil field", *International Transactions in Operational Research*. Vol. 4, No. 2, pp. 125-137.
- Cox, J. C., J. E. Ingersoll, and S. A. Ross 1985. "A Theory of the Term Structure of Interest Rates," *Econometrica*, Vol. 53(2), pp. 385-408.
- DIAS, JOSÉ CARLOS and JOÃO PEDRO VIDAL NUNES 2011. "PRICING REAL OPTIONS UNDER THE CONSTANT ELASTICITY OF VARIANCE DIFFUSION," *The Journal of Futures Markets*, Vol. 31, No. 3, 230 - 250.

- Dias, M. A. G. and K. M. C. Rocha 1999. "Petroleum concessions with extendible options: Investment timing and value using mean reversion and jump processes for oil prices", *Texto Para DISCUSS* □ O N° 620.
- Dixit, A. K. and R. S. Pindyck 1994. "Investment under uncertainty", Princeton, New Jersey.
- Dothan, U. 1978. "On The Term Structure of Interest Rates," *Journal of Financial Economics*, Vol. 7(1), pp. 229-264.
- Gibson, R. and E. S. Schwartz 1990. "Stochastic convenience yield and the pricing of oil contingent claims," *The Journal of Finance*, Vol. 45, No. 3, pp. 959-976.
- Heston, S. L. 1993. "A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options," *Review of Financial Studies*, 6(2): 327-343.
- Kyriakou, Ioannis, Pouliasis, Panos K., and Nikos C. Papapostolou 2016. "Jumps and stochastic volatility in crude oil prices and advances in average option pricing," *Quantitative Finance*, 16:12, pp. 1859-1873.
- Larsson, K. and Nossman, M. 2011. "Jumps and stochastic volatility in oil prices: Time series evidence", *Energy Economics*, Vol. 33, pp. 504-514.
- Liu, Yue, Sun, Huaping, Zhang, Jijian, and Farhad Taghizadeh-Hesary 2020. "Detection of volatility regime-switching for crude oil price modeling and forecasting," *Resources Policy* 69, pp 1-10.
- Lo, A. W. 1988. "Maximum Likelihood Estimation of Generalized Ito Processes with Discretely Sampled Data," *Econometric Theory*, Vol. 4, pp. 231-247.
- Merton, R. C. 1973. "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, Vol. 4, pp. 141-183.
- Nowman, K. B. and H. Wang 2001. "Modelling commodity prices using continuous time models", *Applied Economics Letters*, Vol. 8, No. 5, pp. 341-345.
- Paddock, J. L., Siegel, D. R. and J. L. Smith 1988. "Option valuation of claims on real assets: The case of offshore petroleum leases", *The Quarterly Journal of Economics*,

Vol. 103, No. 3, pp. 479-508.

Postali, F. A. S. and P. Picchetti 2006. "Geometric Brownian motion and structural breaks in oil prices: A quantitative analysis", *Energy Economics*, Vol. 28, pp. 506-522.

Schwartz, E. S. 1997. "The stochastic behavior of commodity prices: implications for valuation and hedging", *Journal of Finance*, Vol. 52, No. 3, pp. 923-973.

Schwartz, E. S. and J. E. Smith 2000. "Short-term variations and long-term dynamics in commodity prices", *Management Science*, Vol. 46, No. 7, pp. 893-911.

Schwarz, G. 1978. "Estimating the Dimension of a Model," *Annals of Statistics*, Vol. 6(2), pp. 461-464.

Smith, J. E. and K. F. McCardle 1999. "Options in the real world: Lessons learned in evaluating oil and gas investments", *Operational Research*, Vol. 47, pp. 1-15.

Yan, X. 2002. "Valuation of commodity derivatives in a new multi-factor model", *Review of Derivatives Research*, Vol. 5, pp. 251-271.

Vasicek, O. 1977. "An Equilibrium Characterization of the Term Structure," *Journal of Financial Economics*, Vol. 5(2), pp. 177-188.

Yu, J. 2007. "Closed-Form Likelihood Approximation and Estimation of Jump-diffusions with an Application to the Realignment Risk of the Chinese Yuan," *Journal of Econometrics*, Vol. 141, pp. 1245-1280.

ABSTRACT

Estimation of Diffusion Models  
for International Crude Oil Prices

Seungmoon Choi\*

The aim of this study is to find a diffusion that can explain the dynamics of international crude oil price through empirical analysis. To this end, we apply daily data on three primary benchmark crude oil prices of Dubai, Brent, and WTI to the GD-GV model that encompasses most of diffusion models suggested in the literature and GD-CEV model, CKLS model, CIR model and Vasicek model which are also nested by the GD-GV model. We estimate these diffusion models by maximum likelihood estimation method for which log-transition probability density function is approximated by the method developed by Ait-Sahalia (2008). Estimation results are quite similar for all of three crude oil prices. We found that CKLS model best explained the movement of oil prices. Moreover, volatility function plays a more important role in describing the dynamics of oil prices than the drift function. We found some evidence that the price of oil tends to revert to its long-run mean value only when the oil price is low. The elasticity of volatility with respect to oil price is estimated to be about 0.75.

Key Words : Crude Oil Price, Diffusion Process, Maximum Likelihood Estimation

---

\* Professor, School of Economics, University of Seoul. schoi22@uos.ac.kr